

# Multiechelon Lot Sizing with Intermediate Demands and Production Capacity

Ming Zhao

Department of Decision and Information Sciences, C.T. Bauer College of Business, University of Houston, Houston, TX 77204  
mzhao@bauer.uh.edu

Minjiao Zhang

Department of Economics, Finance, and Quantitative Analysis, Kennesaw State University, Kennesaw, GA 30144  
mzhang16@kennesaw.edu

We study a multiechelon supply chain model that consists of a production level and several transportation levels, where the demands can exist in the production echelon as well as any transportation echelons. With the presence of stationary production capacity and general cost functions, our model integrates production, inventory and transportation decisions and generalizes existing literature on many multiechelon lot-sizing models. We first prove that the multiechelon lot sizing with intermediate demands (MLS) is NP-hard, which can also be seen as a single source network flow problem in a two-dimensional grid. For uncapacitated cases, we propose an algorithm to solve the MLS with general concave costs. The algorithm is of polynomial time when the number of echelons with demands is fixed, regardless of at which echelon the demands occur. With fixed-charge costs, an innovative algorithm is developed, which outperforms known algorithms for different variants of MLS and gives a tight, compact extended formulation with much less variables and constraints. For cases with stationary production capacity, we propose efficient dynamic programming based algorithms to solve the problem with general concave costs as well as the fixed-charge transportation costs without speculative motives. More importantly, our algorithms improve the computational complexities of many lot-sizing models with demand occurring at final echelon only, which are the special cases of our MLS model.

*Key words* : lot sizing; dynamic programming; network flow; polynomial-time algorithms; extended formulation

---

## 1. Introduction

The multiechelon supply chain is a multifaceted structure, focusing on the integration of production, distribution and inventory. Contemporary research in this area has provided substantial evidence that integrating these decisions can lead to dramatic increases in efficiency. Consider a serial supply

chain for the production and distribution of a product. Production takes place at a manufacturer, and the items are stored initially at the plants, next distributed to national or regional warehouses, then to the distribution centers and so on. Products are either held as inventory or transported to the next echelon, and are eventually shipped to the retail outlets. Such serial supply chain model can be viewed as a multiechelon lot-sizing problem. As one of the most widely studied problems in production planning, the multiechelon lot-sizing problem has been considered primarily under the assumption that demands only occur at the final echelon.

In this paper, we study a multiechelon lot-sizing problem in series with demands at intermediate echelons, which is of particular importance in supply chain systems with both end-product demand and spare parts (or intermediate products) demand. For example, given the complexity and scope of its operations, Ford Motor Company has up to ten echelons of suppliers between itself and its raw materials. Its first-echelon suppliers include 1,400 companies across 4,400 manufacturing sites (see Simchi-Levi et al. (2015)). Besides assembling finished vehicles to satisfy demands of final products, Ford also provides spare vehicle body components for repair shops and dealerships, which are intermediate demands at its upper echelons. A multiechelon supply chain could involve several companies. In such case, different customer channels of different companies have to be considered through intermediate demands in the supply chain. Consider a retailer (e.g., Microsoft, Apple) that ships products from its distribution centers directly to the customers who order through an official website, or to their retail outlets (e.g., Best Buy), as a supply chain that involves a multichannel managed by different companies (see Niranjana and Ciarallo (2011)). In a value-added production system, a product needs to be transported through a sequence of production facilities. The intermediate goods created in this production system usually have their own demands. It is important to fulfill the demand for both the intermediate products and final products.

A generic multiechelon lot-sizing problem that describes all the examples above consists of a production level and several transportation levels, and a level where the final product is sold to the customers. Distinct from much of the existing literature, the demands can exist in the production echelon as well as any transportation echelon. The goal is to determine the production and transportation plan over a finite horizon to meet the demands, which are dynamic and time varying in each echelon, at a minimal total cost. Integrating production, transportation and inventory in the supply chain is crucial when costs are nonlinear, e.g., exhibit economies of scale, and resources are limited. The cost functions are usually assumed to be fixed-charge or general concave functions. In practice, regardless of production level or transshipment level, capacities need to be imposed at each echelon. Following most of the literature on tractable cases, in this paper, we consider a stationary capacity at the production echelon only. Before giving the literature review, we first present the mathematical formulation of the multiechelon lot-sizing problem with intermediate demands.

### 1.1. Mathematical model and notations

Let  $T$  be the length of the planning horizon, and  $L$  be the number of echelons in a serial supply chain, where the manufacturing occurs at the first echelon and products are transported from one echelon to the next echelon to satisfy demands. We denote  $\mathcal{L} = \{1, \dots, L\}$  as the set of all echelons and  $\mathcal{T} = \{1, \dots, T\}$  as the set of all time periods. For each echelon  $i \in \mathcal{L}$  and period  $t \in \mathcal{T}$ , we define the following notations:

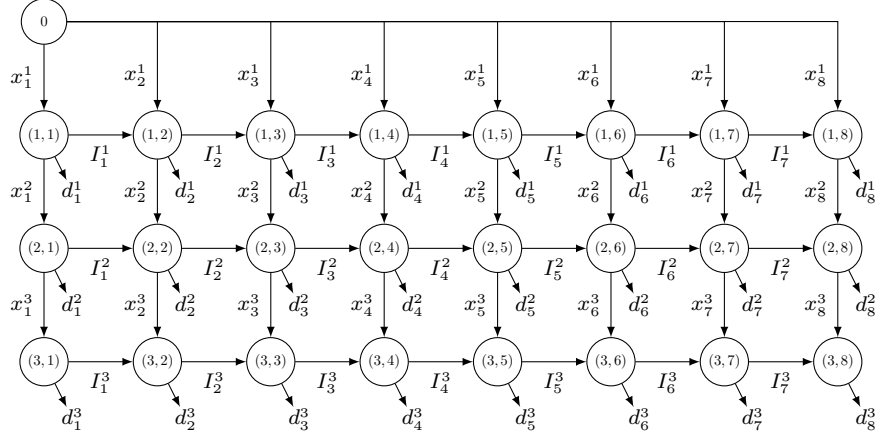


Figure 1 Eight-period three-echelon lot-sizing network

- $d_t^i$ : demand faced by the customer at echelon  $i$  in period  $t$ .
- $\mathbb{C}$ : constant production capacity at the production echelon, i.e., the first echelon.
- $x_t^i$ : production or transportation quantity in period  $t$  at echelon  $i$ . If  $i = 1$ , it is the production quantity, otherwise it is the transportation quantity from echelon  $i - 1$  to  $i$ .
- $p_t^i(x_t^i)$ : production or transportation cost function at echelon  $i$  in period  $t$  for nonnegative amount  $x_t^i$ . In this paper, we will mainly focus on concave cost functions, except for the section 3.3 where we study the outbound transportation with stepwise and non-concave transportation cost functions (see Lee et al. (2003)). The production cost with a fixed charge, as one of the most important cost structures, has received substantial attention in literature. In such a case, we have  $p_t^i(x_t^i) = f_t^i \delta(x_t^i) + c_t^i x_t^i$  where  $f_t^i$  is the fixed cost,  $c_t^i$  is production or transportation cost and  $\delta(x)$  is an indicator function taking the value 1 if  $x > 0$  and 0 otherwise.
- $I_t^i$ : inventory quantity held at echelon  $i$  at the end of period  $t$ .
- $h_t^i(I_t^i)$ : concave inventory holding cost function at echelon  $i$  for nonnegative amount  $I_t^i$  at the end of period  $t$ . A widely used holding cost function is a linear function that we simply denote  $h_t^i(I_t^i) = h_t^i \cdot I_t^i$ .

As an example, the network flows of an eight-period three-echelon lot-sizing problem is shown in Figure 1. Throughout this paper, for an echelon  $i$ , we refer to echelon  $j$  as a *lower* echelon (or level) if  $j > i$  and a *higher* echelon (or level) if  $j < i$  based on their positions in the network, see Figure 1. We let  $[i, j]$  denote the interval  $\{i, i + 1, \dots, j\}$  for  $i \leq j$ , and  $[i, j] = \emptyset$  for  $i > j$ . Given the notations defined above, in a capacitated multiechelon lot-sizing problem with intermediate demands (MCLS), we can minimize the total cost of the production and transportation plan through a serial structure supply chain as follows,

$$\min \sum_{i=1}^L \sum_{t=1}^T (p_t^i(x_t^i) + h_t^i(I_t^i)) \quad (1a)$$

$$\text{s.t. } I_{t-1}^i + x_t^i = I_t^i + x_t^{i+1} + d_t^i \quad \forall i \in \mathcal{L} \setminus L, t \in \mathcal{T} \quad (1b)$$

$$I_{t-1}^L + x_t^L = I_t^L + d_t^L \quad \forall t \in \mathcal{T} \quad (1c)$$

$$x_t^1 \leq \mathbb{C} \quad \forall t \in \mathcal{T} \quad (1d)$$

$$I_0^i = I_T^i = 0 \quad \forall i \in \mathcal{L} \quad (1e)$$

$$x_t^i \geq 0, I_t^i \geq 0 \quad \forall i \in \mathcal{L}, t \in \mathcal{T} \quad (1f)$$

As one of the important cost structures, the case with fixed-charge production and transportation costs and linear holding costs is widely used in practice and often appears in the literature. In this paper, we call it as *the fixed-charge cost structure* for short.

Note that we have an uncapacitated multiechelon lot-sizing problem with intermediate demands (MULS) if (1d) is dropped. When the number of echelons  $L$  is given, we denote the corresponding problems as  $L$ -ULS or  $L$ -CLS for uncapacitated or capacitated cases respectively. For the cases that demand occurs only at the final echelon, we append “-F” to the abbreviations. For example, we use MULS-F to indicate the uncapacitated multiechelon lot-sizing problem with demands occurring at the final echelon only. Similarly, we have abbreviations MCLS-F,  $L$ -ULS-F and  $L$ -CLS-F. For notational convenience, we define  $d^i(s, t) = \sum_{j=s}^t d_j^i$ , i.e., the cumulative demand at echelon  $i$  in periods from  $s$  to  $t$ , and it is 0 if  $s > t$ . To ensure the feasibility of the MCLS, we assume that  $\sum_{\ell=1}^L d^\ell(1, t) \leq t\mathbb{C}$ .

Different from the lot-sizing problems with demands occurring at the final echelon only, we need to characterize demand satisfaction status at each echelon, which motivates us to introduce  $L$ -dimensional vectors. So we define a set  $\mathcal{V} = \{v \in [0, T]^L : v_1 \leq \dots \leq v_L\}$ . Let  $\mathcal{L}_0$  be a subset of the set  $\mathcal{L}$  such that any echelon  $i \in \mathcal{L}_0$  is just a transshipment echelon without intermediate demands,  $\mathcal{L}_1 = \mathcal{L} \setminus \mathcal{L}_0$  and  $L_k = |\mathcal{L}_k|$  for  $k = 0, 1$ . Without loss of generality, we can assume that the last echelon  $L \in \mathcal{L}_1$ . We also define a relaxed set of  $\mathcal{V}$  by fixing the components corresponding to the transshipment echelons to 0, i.e.,  $\bar{\mathcal{V}} = \{v \in [0, T]^L : v_i \leq v_j \forall i \leq j \in \mathcal{L}_1 \text{ and } v_i = 0 \forall i \in \mathcal{L}_0\}$ . For notational convenience, we denote  $\mathbf{0} \in \mathcal{V}$  with  $\mathbf{0}_i = 0 \forall i \in \mathcal{L}$ ,  $\mathbf{1} \in \mathcal{V}$  with  $\mathbf{1}_i = 1 \forall i \in \mathcal{L}$ ,  $\mathbf{T} \in \mathcal{V}$  with  $\mathbf{T}_i = T \forall i \in \mathcal{L}$ ,  $\bar{\mathbf{1}} \in \bar{\mathcal{V}}$  with  $\bar{\mathbf{1}}_i = 1 \forall i \in \mathcal{L}_1$  and  $\bar{\mathbf{T}} \in \bar{\mathcal{V}}$  with  $\bar{\mathbf{T}}_i = T \forall i \in \mathcal{L}_1$ .

## 1.2. Literature review

The study of lot-sizing problems starts from the seminal paper Wagner and Whitin (1958), in which they propose an  $O(T^2)$  algorithm to solve the uncapacitated single-echelon lot-sizing problem (1-ULS) based on the properties of the extreme solutions. Decades later, the algorithm to the 1-ULS problem was improved by Federgruen and Tzur (1991), Aggarwal and Park (1993), and Wagelmans et al. (1992) to  $O(T \log T)$ . The single-echelon lot-sizing problem with varying capacity is known to be NP-hard, see Bitran and Yanasse (1982), Florian et al. (1980). However, by considering a constant capacity over the planning horizon (i.e., 1-CLS), Florian and Klein (1971) develop an  $O(T^4)$  dynamic programming algorithm to solve the problem, which was improved to  $O(T^3)$  by van Hoesel and Wagelmans (1996) when the holding costs are linear. An interesting extension of constant capacity is considering constant batch size. When backlogging is not allowed, Pochet and Wolsey (1993) give a  $O(T^3)$  algorithm. With backlogging, Van Vyve (2007) also gives an  $O(T^3)$  algorithm for a general number of maximum batches. A detailed study of many lot-sizing models can be found in Pochet and Wolsey (2006).

Uncapacitated multiechelon lot-sizing problem is first studied by Zangwill (1969) with demand occurring at the final echelon only. The MULS-F is modeled as a network flow in a two-dimensional grid and an  $O(LT^4)$  dynamic programming algorithm is proposed by Zangwill (1969). Later, Love (1972) gives an  $O(LT^3)$  algorithm by exploiting a nested structure based on the assumption that the production costs are nonincreasing over time and the holding costs are nondecreasing over echelons. Atamtürk and Küçükyavuz (2008) develop an  $O(T^2)$  algorithm for 1-ULS with inventory upper bounds and the polyhedral structure is studied in Atamtürk and Küçükyavuz (2005). Because of its importance in applications, 2-ULS-F receives a lot of attention. van Hoesel et al. (2005) show that 2-ULS-F can be solved in  $O(T^3)$  time, which is improved to  $O(T^2 \log T)$  by Melo and Wolsey (2010). Lee et al. (2003) study an application of 2-ULS-F with backlogging allowed in the last echelon and outbound shipment. This results in a stepwise and non-concave transportation cost for the shipment between the two echelons. Lee et al. (2003) show that the problem can be solved in  $O(T^6)$  time. Hwang (2010) generalizes the model by considering general concave production costs. To the best of our knowledge, Zhang et al. (2012) are the first that study the complexity of multiechelon lot-sizing problem with intermediate demands, more specifically 2-ULS. Zhang et al. (2012) show that 2-ULS can be solved in  $O(T^4)$  time with a dynamic programming algorithm, which implies an extended formulation with  $O(T^4)$  variables and  $O(T^4)$  constraints.

Kaminsky and Simchi-Levi (2003) study a three-echelon lot-sizing model with capacities, which can be reduced to a two-echelon model. With linear holding costs and no speculative motives assumption, the model can be solved in  $O(T^8)$  time, where no speculative motives implies that

it is optimal to transport to the lower echelons as late as possible, i.e.,  $c_t^\ell + h_t^{\ell+1} \geq c_{t+1}^\ell + h_t^\ell$  in the case of the fixed-charge cost structure. van Hoesel et al. (2005) provide a detailed analysis of capacitated multiechelon lot-sizing problem MCLS-F. They show that, in the case of fixed-charge transportation costs without speculative motives, the MCLS-F can be solved in polynomial time  $O(T^7 + LT^4)$  and the algorithm complexity can be improved to  $O(T^6)$  when  $L = 2$ . However, they only provide a pseudo-polynomial algorithm with complexity  $O(LT^{2L+3})$  for MCLS-F with general concave costs. Later, this result is largely improved by Hwang et al. (2013) as the MCLS-F is proved to be polynomial solvable in  $O(LT^8)$  time by using a new concept called basis paths.

It is well-known that the lot-sizing problem with concave costs can be seen as a minimal concave cost network flow problem in a two-dimensional grid with only one source. As much literature on network flow problems focuses on more general setting where the network could have multiple sources, we only find two papers, by He et al. (2015) and Ahmed et al. (2016), that are very related to our work. For more research on general network flow problem, we refer to the literature in those two papers. He et al. (2015) focus on the grid network with multiple sources at the first level. They show that MULS, as a special case, is polynomial solvable when the number of echelons  $L$  is fixed but the computational complexity is not specified. Ahmed et al. (2016) study a similar grid network with flow capacities and classify many NP-hard and polynomial solvable cases. They show that the minimal concave cost network flow problem in a two-dimensional grid with two sources at top level is NP-hard. However, the complexity of solving the single source problem, MULS, is unknown.

### 1.3. Main contributions and outline

In this paper, our focus is on studying the computational complexity and developing efficient polynomial time algorithms for solving the multiechelon lot-sizing problem with intermediate demands. Our results generalize many previous studies. Because of the intermediate demands, the problem considered in this paper more practically formulates actual multiechelon systems and is thus more widely applicable to actual inventory and supply networks.

Our main contributions are presented in each of the following sections in details and listed as follows,

- In Section 2, we show that the MULS with the fixed-charge cost structure is NP-hard. As this is a classical single source network flow problem, we establish an important computational complexity boundary.
- In Sections 3 and 4, we provide efficient algorithms for solving MULS and MCLS respectively (see Table 1). For the uncapacitated cases, we:

- generalize the results of Zangwill (1969) to solve the MULS with general concave costs (Subsection 3.1);
- propose efficient dynamic programming recursions based on shortest path algorithm to solve the MULS with the fixed-charge cost structure, and a tight extended formulation with much less variables and constraints than the one suggested by Zhang et al. (2012) (Subsection 3.2); and
- extend the dynamic programming recursions in the previous subsection to solve 2-ULS with backlogging and outbound transportation where the transportation cost is stepwise and non-concave (Subsection 3.3).

For the capacitated cases, we show that all possible values of cumulative production and transportation quantities can be enumerated in polynomial time efficiently. Then, following the approaches by Florian and Klein (1971) and van Hoesel et al. (2005), the application of all allowable values of cumulative production and transportation quantities gives dynamic programming recursions to solve MCLS with

- general concave costs (Subsection 4.3), and
- the fixed-charge transportation cost without speculative motives (Subsection 4.4).

The computational complexities of all the studied models in this paper are presented in Table 1. It is important to mention that some models are polynomial solvable when  $L_1$  is fixed. Thus, the polynomial solvability depends on the number of echelons with demands instead of the levels they reside in the supply chain.

- Importantly, we achieve improved complexities in solving several lot-sizing models comparing the best known algorithms proposed in literature. The comparisons are shown in Table 2. Except the last row in Table 2, all comparisons are for the cases with demand occurring at the final echelon only, because our algorithms are general and those models can be solved as special cases.

At last, we conclude our paper in Section 5.

## 2. NP-hardness

Much effort has been spent on finding efficient polynomial-time algorithms or proof of NP-hardness on the dynamic lot-sizing problem and its extensions to certain network flow problems. On the one hand, Zangwill (1969) show that if the demand occurs at the final echelon only, then MULS is polynomial solvable with a complexity of  $O(LT^4)$ . On the other hand, Ahmed et al. (2016) show that the minimum concave cost flow in a two-dimensional grid is NP-hard if there are two sources at the first level. An obvious gap between these results is the unknown complexity of the MULS, which can also be viewed as a single source network flow problem in a two-dimensional grid. The main result of this section, Theorem 1, is to close this gap by showing that:

Model	Cost structure	Complexity
MULS	general concave costs	$O(LT^{3L_1+1})$
	fixed-charge cost structure	$O(L \min(T^L \log T, T^{3L_1+1}))$
2-ULS <sup>a</sup>	stepwise and non-concave transportation costs	$O(T^5)$
MCLS	general concave costs	$O(LT^{2L_1L+2})$
	fixed-charge transportation costs and no speculative motives	$O(T^{4L_1+2} + LT^{3L_1+1})$

<sup>a</sup> with backlogging and outbound transportation

**Table 1** Computational complexity results.

Model	The best known results		Results of this paper
3-ULS-F <sup>b</sup>	$O(T^4)$	by Zangwill (1969)	$O(T^3 \log T)$
2-ULS-F <sup>c</sup>	$O(T^6)$	by Lee et al. (2003)	$O(T^5)$
MCLS-F <sup>a</sup>	$O(LT^{2L+3})$	by van Hoesel et al. (2005)	$O(LT^{2L+2})$
	$O(LT^8)$	by Hwang et al. (2013)	
MCLS-F <sup>d</sup>	$O(T^7 + LT^4)$	by van Hoesel et al. (2005)	$O(T^6 + LT^4)$
2-CLS-F <sup>d</sup>	$O(T^6)$	by van Hoesel et al. (2005)	$O(T^5)$
2-ULS <sup>b</sup>	$O(T^4)$	by Zhang et al. (2012)	$O(T^2 \log T)$

<sup>a</sup> with general concave cost functions

<sup>b</sup> with fixed-charge cost structure

<sup>c</sup> with backlogging and outbound transportation

<sup>d</sup> with fixed-charge transportation costs and no speculative motives

**Table 2** Computational complexity comparisons between our results and the best known results.

**THEOREM 1.** *The MULS with the fixed-charge cost structure is NP-hard.*

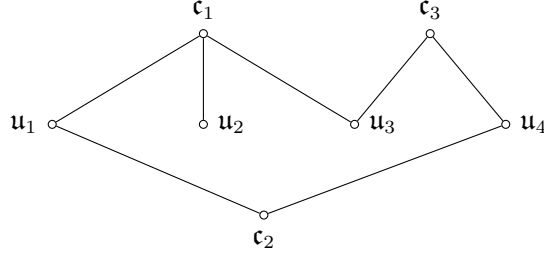
The proof of Theorem 1 is a reduction from planar 3SAT, which is a particular version of the well-known satisfiability problem 3SAT and is proved to be NP-complete by Lichtenstein (1982):

**Instance:** Consider a set of variables  $\mathcal{U}$  and a set of clauses  $\mathcal{C}$ . Each clause  $\mathbf{c}_j \in \mathcal{C}$  contains at most three literals, where a literal is either a variable  $\mathbf{u}_i \in \mathcal{U}$  or the negation of a variable. Furthermore, the identification graph  $\mathfrak{G} = (\mathcal{U} \cup \mathcal{C}, \mathbb{E})$  is planar, where  $\mathbb{E} = \{(\mathbf{u}_i, \mathbf{c}_j) | \mathbf{u}_i \in \mathbf{c}_j \text{ or } \bar{\mathbf{u}}_i \in \mathbf{c}_j\}$ .

**Questions:** Is there an assignment for the variables such that all clauses are satisfied?

An example of planar 3SAT is illustrated in Figure 2.

Next, we give an overall strategy of the proof and refer the reader to Appendix A for more details.



**Figure 2** Graph  $\mathfrak{G}$  of a planar 3SAT instance with  $c_1 = u_1 \vee u_2 \vee \bar{u}_3$ ,  $c_2 = \bar{u}_1 \vee \bar{u}_4$  and  $c_3 = \bar{u}_3 \vee u_4$ .

**Overall strategy** Given the identification graph  $\mathfrak{G}$  of a planar 3SAT instance, the overall strategy of proving Theorem 1 includes three steps. In the first step, an embedding procedure with slight modifications over the one developed by Shi and Su (2015) is applied to the planar graph  $\mathfrak{G}$ . The resulting graph  $\mathfrak{R}$  is a subgraph of a two-dimensional grid network satisfying certain structural conditions. In the second step, a basic flow network with four source nodes and four demand nodes (see Figure EC.5 in Appendix A) is designed to replace each vertex of the embedded graph  $\mathfrak{R}$ . Interestingly, the basic flow network has only two optimal solutions so that each solution can represent either 0 or 1, i.e., the value of a variable in  $\mathcal{U}$  or its negation. We connect those basic flow networks together in a fashion to obtain a single source network in a two-dimensional grid and achieve minimal cost when all clauses have values of 1. In the final step, Theorem 1 is proved by showing that the minimal cost of the MULS problem can identify whether the given planar 3SAT instance has a satisfying solution.

Theorem 1 demonstrates that the MULS is very different from MULS-F and MCLS-F in computational complexity since even the latter one is proved to be polynomial solvable by Hwang et al. (2013). In the following sections, we focus on developing efficient polynomial algorithms for MULS and MCLS given that the number of echelons is fixed.

### 3. Uncapacitated cases

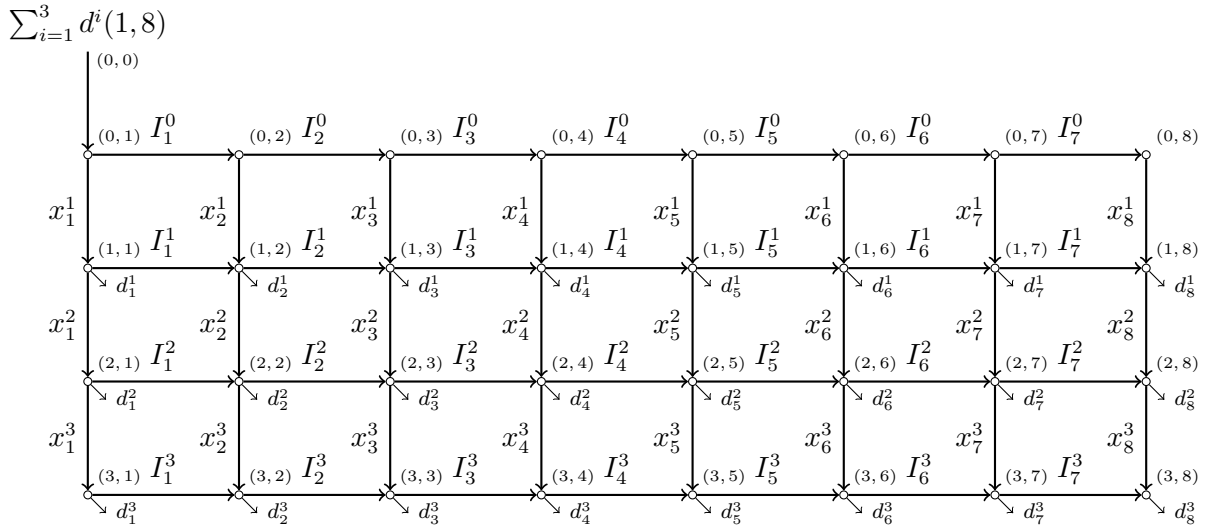
In this section, we first show that MULS with general concave costs can be solved in  $O(LT^{3L_1+1})$  time by generalizing Zangwill (1969)’s approach. It is important to note that the problem is polynomial solvable when  $L_1$ , as opposed to  $L$ , is fixed. Then, we propose a shortest path algorithm on a network with node set  $\mathcal{V}$  to solve MULS with the fixed-charge cost structure, which is quite different from the traditional approach on the network with nodes of echelon and period pairs (such as Figure 1). Importantly, the algorithm can be further improved by using dynamic programming recursions and outperforms many best known algorithms in literature. At last, we show that this approach can be extended to solve 2-ULS with backlogging and outbound transportation where the transportation cost functions are stepwise and non-concave.

### 3.1. General concave costs

Following the traditional view by Zangwill (1969), the MULS with concave cost functions can be seen as a minimal concave flow problem in a two-dimensional grid. Figure 3 shows the network flow model of a 3-ULS. The node  $(i, t)$  represents the time period  $t$  at echelon  $i$ . We add an echelon at level 0 to aggregate all productions into a single source. The source node  $(0, 0)$  has  $\sum_{i=1}^L d^i(1, T)$  units of products available, which must be sent through the grid network to supply all demands. Note that  $I_{0,t} \forall t \in [1, T]$  are introduced for notational convenience and they have no cost. We denote  $\bar{v}_i \in \bar{\mathcal{V}}$  with all components 0 except the  $i$ -th component is 1 if  $i \in \mathcal{L}_1$ , and

$$\mathcal{D}_i(\bar{v}, \bar{w}) = \sum_{\ell=i}^L d^\ell(\bar{v}_\ell + 1, \bar{w}_\ell).$$

for  $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$  and  $i \in [1, L]$ .



**Figure 3** Network flow of 3-ULS on a two-dimensional grid.

When demand occurs at the final echelon only, Zangwill (1969) shows that the total units shipped into a given node  $(i, t)$  are demands of a sequence of consecutive periods after period  $t - 1$  at the final echelon, e.g.,  $d^L(a, b)$  with  $t \leq a \leq b \in [1, T]$ . But, it is far more complicated to describe the shipment amount sent into a node when intermediate demands exist. Given a node  $(i, t)$  and  $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$  with  $\bar{v}_\ell \leq \bar{w}_\ell \forall \ell \in [i, L]$ , we define  $F_{i,t}(\bar{v}, \bar{w})$  as the minimal total cost of shipping units of  $\mathcal{D}_i(\bar{v}, \bar{w})$  into node  $(i, t)$  to satisfy all the demands of destination nodes  $(\ell, \bar{v}_\ell + 1), \dots, (\ell, \bar{w}) \forall \ell \in [i, L]$ . However, the choices of  $\bar{v}, \bar{w}$  are not arbitrary and we have the following observation.

**Observation 1**  $F_{i,t}(\bar{v}, \bar{w})$  is well defined if

1.  $(\bar{v}, \bar{w})$  is a legitimate pair from level  $i$  such that, for any  $j \in [i, L-1] \cap \mathcal{L}_1$ , we have either  $\bar{v}_{j+1} < \bar{w}_j$  or  $\bar{v}_\ell = \bar{w}_\ell \forall \ell \in [j+1, L]$ .
2.  $(\bar{v}, \bar{w})$  satisfies conditions Feas $_{i,t-1}$  and 2 as follows
  - Feas $_{i,t-1}$ :  $\bar{v}_i = t - 1 < \bar{w}_i$  when  $i \in \mathcal{L}_1$ ;
  - Feas $_{i,t-2}$ :  $\bar{v}_\ell \geq t - 1 \forall \ell \in [i, L] \cap \mathcal{L}_1$ ;

The arborescent structure of an extreme solution implies that the demand shipped into node  $(i, t)$  is  $\sum_{\ell=i}^L d^\ell(\underline{t}_\ell + 1, \bar{t}_\ell)$  with  $\underline{t}_\ell, \bar{t}_\ell \forall \ell \in [i, L]$  such that  $\underline{t}_{\ell-1} \leq \underline{t}_\ell$  and  $\bar{t}_{\ell-1} \leq \bar{t}_\ell \forall \ell \in [i+1, L]$ . Thus, we can denote the demand units as  $\mathcal{D}_i(\bar{v}, \bar{w})$  for some  $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$  with  $\bar{v}_\ell = \underline{t}_\ell$  and  $\bar{w}_\ell = \bar{t}_\ell \forall \ell \in [i, L] \cap \mathcal{L}_1$  because the echelons in  $\mathcal{L}_0$  have no demand. Since the values of  $\bar{v}_\ell, \bar{w}_\ell \forall \ell \in [1, i-1]$  are irrelevant, we set  $\bar{v}_\ell = \bar{w}_\ell \forall \ell \in [1, i]$ . Note that if  $\underline{t}_{j+1} \geq \bar{t}_j$  for some  $j \in [i, L-1]$ , then the shipment into node  $(i, t)$  cannot serve any demands at echelon  $j+1, \dots, L$  because of the arborescent structure of an extreme solution. So  $\underline{t}_\ell = \bar{t}_\ell$ , i.e.,  $\bar{v}_\ell = \bar{w}_\ell \forall \ell \in [j+1, L]$ . Therefore, we conclude that  $(\bar{v}, \bar{w})$  is a legitimate pair from level  $i$ .

The definition of legitimate pair in Observation 1 is to rule out infeasible cases. For example, it is impossible to have the shipment into node  $(1, 4)$  fulfill demand  $\mathcal{D}_1((3, 5, 6), (5, 7, 8)) = d^1(4, 5) + d^2(6, 7) + d^3(7, 8)$ , because the shipment cannot reach node  $(2, 6)$  or  $(2, 7)$  from node  $(1, 4)$  or  $(1, 5)$  at level 1 without passing node  $(1, 6)$  or  $(2, 5)$  whose demands however are not specified in  $\mathcal{D}_1((3, 5, 6), (5, 7, 8))$ . So, the definition of the legitimate pair indicates that either the shipment reaches level  $j+1$  (i.e.,  $\bar{v}_{j+1} < \bar{w}_j$ ) or stops at level  $j$  (i.e.,  $\bar{v}_\ell = \bar{w}_\ell \forall \ell \in [j+1, L]$ ). Note that if  $(\bar{v}, \bar{w})$  is a legitimate pair from level  $i$ , then we have  $(\bar{v}, \bar{w})$  as a legitimate pair from level  $i+1$  and  $(\bar{v} + \bar{e}_i, \bar{w})$  as a legitimate pair from level  $i$  by the definition.

In addition, the condition Feas $_{i,t-1}$  in Observation 1 is needed because the demand of node  $(i, t)$  has to be satisfied by the shipment into node  $(i, t)$  if  $i \in \mathcal{L}_1$ . The condition Feas $_{i,t-2}$  states that the shipment into node  $(i, t)$  can only cover demands in or after period  $t$ .

Our goal is to determine the value of  $F_{0,t}(\mathbf{0}, \bar{\mathbf{T}})$ , which is well defined by easily checking the conditions in Observation 1. Now, with a given node  $(i, t)$ , it is important to note that the units shipped into node  $(i, t)$  will be fully distributed to nodes  $(i+1, t)$  and  $(i, t+1)$ . In the rest of this subsection, we suppose that  $F_{i,t}(\bar{v}, \bar{w})$  is well defined with a given node  $(i, t)$  and  $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$ ,  $\bar{v}_\ell \leq \bar{w}_\ell \forall \ell \in [i, L]$ . Without loss of generality, we assume  $\bar{v}_\ell \geq t$  if  $\bar{v}_\ell = \bar{w}_\ell$  with some  $\ell \in [i, L] \cap \mathcal{L}_1$  for notational convenience. It is easy to see that there is no change on the definition of  $F_{i,t}(\bar{v}, \bar{w})$  because  $\bar{v}_\ell = \bar{w}_\ell$  implies  $[\bar{v}_\ell + 1, \bar{w}] = \emptyset$ . Let  $\Omega_{i,t}(\bar{v}, \bar{w})$  be the set of  $\bar{u} \in \bar{\mathcal{V}}$  such that  $\bar{u}_\ell = \bar{v}_\ell \forall \ell \in [1, i]$ ,  $\max(t, \bar{v}_\ell) \leq \bar{u}_\ell \leq \bar{w}_\ell \forall \ell \in [i+1, L] \cap \mathcal{L}_1$ ,  $(\bar{v}, \bar{u})$  is a legitimate pair from level  $i+1$  and  $(\bar{u}, \bar{w})$  is a legitimate pair from level  $i$ . We point out that  $\Omega_{i,t}(\bar{v}, \bar{w}) \neq \emptyset$  because we can set  $\bar{u}_\ell = \bar{v}_\ell \forall \ell \in [1, i]$  and

$\bar{u}_\ell = \bar{w}_\ell \forall \ell \in [i+1, L]$ . The purpose of having the set  $\Omega_{i,t}(\bar{v}, \bar{w})$  is to make sure that the shipments from node  $(i, t)$  to nodes  $(i+1, t)$  and  $(i, t+1)$  can be denoted as  $\mathcal{D}_{i+1}(\bar{v}, \bar{u})$  and  $\mathcal{D}_i(\bar{u} + \bar{e}_i, \bar{w})$  respectively. By introducing two conditions (C1)-(C2),

$$(C1): \quad \bar{v}_{i+1} \geq t \text{ and } i+1 \in \mathcal{L}_1;$$

$$(C2): \quad \bar{v}_i + 1 = \bar{w}_i \text{ and } i \in \mathcal{L}_1.$$

we can generalize one of the main theorems by Zangwill (1969) as in Proposition 1. The conditions (C1) and (C2) identify the cases that the shipment sending out from node  $(i, t)$  is sent to only one node  $(i+1, t)$  or  $(i, t+1)$ , which gives boundary conditions of the dynamic programming recursions.

**PROPOSITION 1.** *The shipment pattern of an extreme solution from node  $(i, t)$  can be characterized into four cases.*

1. *If conditions (C1) and (C2) do not hold, and  $i < L, t < T$ , then there exists a vector  $\bar{u} \in \Omega_{i,t}(\bar{v}, \bar{w})$  such that the amount shipped into node  $(i, t)$  is sent to node  $(i+1, t)$  and  $(i, t+1)$  with amount  $\mathcal{D}_{i+1}(\bar{v}, \bar{u})$  and  $\mathcal{D}_i(\bar{u} + \bar{e}_i, \bar{w})$  respectively and*

$$x_t^{i+1} = \mathcal{D}_{i+1}(\bar{v}, \bar{u}) \text{ and } I_t^i = \mathcal{D}_i(\bar{u} + \bar{e}_i, \bar{w}).$$

*In this case, we have the dynamic programming recursion*

$$F_{i,t}(\bar{v}, \bar{w}) = \min_{\bar{u} \in \Omega_{i,t}(\bar{v}, \bar{w})} \left[ p_t^{i+1}(\mathcal{D}_i(\bar{v}, \bar{u})) + F_{i+1,t}(\bar{v}, \bar{u}) + h_t^i(\mathcal{D}_i(\bar{u} + \bar{e}_i, \bar{w})) + F_{i,t+1}(\bar{u} + \bar{e}_i, \bar{w}) \right], \quad (2)$$

*with  $F_{i+1,t}(\bar{v}, \bar{u})$  and  $F_{i,t+1}(\bar{u} + \bar{e}_i, \bar{w})$  to be well defined.*

2. *if condition (C1) holds or  $i = L$ , then the amount shipped into node  $(i, t)$  is fully sent to node  $(i, t+1)$  with  $I_t^i = \mathcal{D}_i(\bar{v} + \bar{e}_i, \bar{w})$ . In this case, we have the boundary condition*

$$F_{i,t}(\bar{v}, \bar{w}) = h_t^i(\mathcal{D}_i(\bar{v} + \bar{e}_i, \bar{w})) + F_{i,t+1}(\bar{v} + \bar{e}_i, \bar{w}),$$

*with  $F_{i,t+1}(\bar{v} + \bar{e}_i, \bar{w})$  to be well defined.*

3. *if condition (C2) holds or  $t = T$ , then the amount shipped into node  $(i, t)$  is fully sent to node  $(i+1, t)$  with  $x_t^{i+1} = \mathcal{D}_i(\bar{v}, \bar{w})$  if  $i < L$ . In this case, we have the boundary condition*

$$F_{i,t}(\bar{v}, \bar{w}) = p_t^{i+1}(\mathcal{D}_i(\bar{v}, \bar{w})) + F_{i+1,t}(\bar{v}, \bar{w}),$$

*with  $F_{i+1,t}(\bar{v}, \bar{w})$  to be well defined.*

4. *if condition (C1)-(C2) hold or  $(i, t) = (L, T)$ , then there is no shipment to be sent out from node  $(i, t)$ . In this case, we have  $F_{i,t}(\bar{v}, \bar{w}) = 0$ .*

*Note that the vector  $\bar{e}_i$  in defining  $I_t^i$  indicates that the demand  $d_t^i$  is consumed at node  $(i, t)$ .*

*Proof* Let us consider the first case where  $i < L, t < T$ . It is well known that the flow corresponding to any extreme solution is acyclic (see, e.g., Ahuja et al. (1993)) in uncapacitated network flow problems. Thus, the total shipment  $\mathcal{D}_i(\bar{v}, \bar{w})$  into node  $(i, t)$  will be sent to nodes  $(i+1, t)$  and  $(i, t+1)$ , with amount  $\sum_{\ell=i+1}^L d^\ell(\bar{v}_\ell + 1, \bar{u}_\ell)$  and  $d^i(t+1, \bar{w}_i) + \sum_{\ell=i+1}^L d^\ell(\bar{u}_\ell + 1, \bar{w}_\ell)$  respectively. Since  $\bar{u}_\ell \forall \ell \leq i$  are not involved in the description, we can set  $\bar{u}_\ell = \bar{v}_\ell \forall \ell \leq i$  for the convenience of recursions. Because of the arborescent structure of an extreme solution, we know  $\bar{v}_\ell \leq \bar{u}_\ell \leq \bar{w}_\ell \forall \ell \in [i+1, L] \cap \mathcal{L}_1$ . Similar to the reason for introducing Observation 1, we will discuss a few more properties about  $\bar{u}_\ell$ s because they are used to describe the amount shipped into nodes  $(i+1, t)$  and  $(i, t+1)$  and cannot be arbitrary. First, we need  $(\bar{v}, \bar{u})$  to be legitimate from level  $i+1$  and  $(\bar{u}, \bar{w})$  to be legitimate from level  $i$ . Next, because the shipment  $d^i(t+1, \bar{w}_i) + \sum_{\ell=i+1}^L d^\ell(\bar{u}_\ell + 1, \bar{w}_\ell)$  into node  $(i, t+1)$  only covers demands after period  $t$ , we have  $\bar{u}_\ell \geq t$  for all  $\ell \in [i+1, L] \cap \mathcal{L}_1$ . By summarizing the above discussions, we conclude that  $\bar{u} \in \bar{\mathcal{V}}$ ,  $\bar{u} \in \Omega_{i,t}(\bar{v}, \bar{w})$  and the shipment sent to node  $(i+1, t)$  (i.e.,  $x_t^{i+1}$ ) and  $(i, t+1)$  (i.e.,  $I_t^i$ ) can be written as  $\mathcal{D}_{i+1}(\bar{v}, \bar{u})$  and  $\mathcal{D}_i(\bar{u} + \bar{e}_i, \bar{w})$  respectively. The vector  $\bar{e}_i$  is used to indicate that the demand  $d_t^i$  is consumed at node  $(i, t)$ . Note that  $d^i(t+1, \bar{w}_i) = d^i(\bar{u}_i + \bar{e}_i, \bar{w}_i)$  because the condition Feas $_{i,t-1}$  implies  $\bar{u}_i = \bar{v}_i = t-1$  if  $i \in \mathcal{L}_1$ . It also implies that  $\bar{u} + \bar{e}_i \in \bar{\mathcal{V}}$  as  $\bar{u}_i = \bar{v}_i = t-1 \leq t \leq \bar{u}_\ell \forall \ell \in [i+1, L] \cap \mathcal{L}_1$  if  $i \in \mathcal{L}_1$ .

With the shipping pattern, the recursion (2) follows immediately by finding  $\bar{u}$  at a minimal total cost. Now, we need to show that  $F_{i+1,t}(\bar{v}, \bar{u})$  and  $F_{i,t+1}(\bar{u} + \bar{e}_i, \bar{w})$  are well defined. We consider  $F_{i+1,t}(\bar{v}, \bar{u})$  first. We have already shown that  $(\bar{v}, \bar{u})$  is legitimate from level  $i+1$ . Since the condition (C1) does not hold, we have  $\bar{v}_{i+1} = t-1$  if  $i+1 \in \mathcal{L}_1$ . Then  $\bar{u} \in \Omega_{i,t}(\bar{v}, \bar{w})$  implies  $t \leq \bar{u}_{i+1}$  and  $\bar{v}_{i+1} = t-1 < t \leq \bar{u}_{i+1}$  if  $i+1 \in \mathcal{L}_1$ . The condition Feas $_{i+1,t-1}$  holds for  $(\bar{v}, \bar{u})$ . The condition Feas $_{i+1,t-2}$  holds obviously for  $(\bar{v}, \bar{u})$  because Feas $_{i,t-2}$  holds for  $(\bar{v}, \bar{w})$ . Then, we consider  $F_{i,t+1}(\bar{u} + \bar{e}_i, \bar{w})$ . That  $(\bar{u}, \bar{w})$  is legitimate from level  $i$  implies that  $(\bar{u} + \bar{e}_i, \bar{w})$  is legitimate from level  $i$  by definition. Since the condition (C2) does not hold, we have  $\bar{v}_i + 1 < \bar{w}_i$  if  $i \in \mathcal{L}_1$ . Then,  $(\bar{u} + \bar{e}_i)_i = \bar{u}_i + 1 = \bar{v}_i + 1 = t < \bar{w}_i$  if  $i \in \mathcal{L}_1$ . The condition Feas $_{i,t+1-1}$  holds for  $(\bar{u} + \bar{e}_i, \bar{w})$ . The condition Feas $_{i,t+1-2}$  holds obviously for  $(\bar{u} + \bar{e}_i, \bar{w})$  because  $\bar{u} \in \Omega_{i,t}(\bar{v}, \bar{w})$  and  $t \leq \bar{u}_\ell \forall \ell \in [i+1, L] \cap \mathcal{L}_1$ .

Next, we consider the final case, i.e., the fourth case. It holds obviously if  $(i, t) = (L, T)$ . So, we suppose that the conditions (C1) and (C2) hold. When the condition (C1) holds, the amount shipped into node  $(i, t)$  is fully sent to node  $(i, t+1)$ , because, if  $i+1 \in \mathcal{L}_1$ ,  $\bar{v}_{i+1} \geq t$  and the demand of node  $(i+1, t)$  is not covered by the shipment into  $(i, t)$ , i.e., no shipment is sent from node  $(i, t)$  to  $(i+1, t)$ . When condition (C2) holds, the amount shipped into node  $(i, t)$  is fully sent to node  $(i+1, t)$  because  $\bar{v}_i = t-1$  (since the condition Feas $_{i,t-1}$  holds for  $(\bar{v}, \bar{w})$ ),  $t = \bar{v}_i + 1 = \bar{w}_i$  and the demand of node  $(i, t+1)$  is not covered by the shipment into  $(i, t)$ , i.e., no shipment sending from node  $(i, t)$  to  $(i, t+1)$ . Since there is no shipment sending out from node  $(i, t)$ , we have boundary condition  $F_{i,t}(\bar{v}, \bar{w}) = 0$ .

Finally, we show the second and third cases to conclude the proof. The second case holds obviously if  $i = L$ . As shown before, the condition (C1) implies that the amount shipped into node  $(i, t)$  is fully sent to node  $(i, t + 1)$ . We only need to show that  $F_{i,t+1}(\bar{v} + \bar{e}_i, \bar{w})$  is well defined. Note that  $(\bar{v} + \bar{e}_i, \bar{w})$  is legitimate from level  $i$  because  $(\bar{v}, \bar{w})$  is legitimate from level  $i$ . We can assume that condition (C2) does not hold, otherwise it becomes the last case, which will be shown later. It implies that  $\bar{w}_i > \bar{v}_i + 1$  if  $i \in \mathcal{L}_1$ . Also that the condition  $\text{Feas}_{i,t-1}$  holds for  $(\bar{v}, \bar{w})$  implies  $(\bar{v} + \bar{e}_i)_\ell \geq t \forall \ell \in [i, L] \cap \mathcal{L}_1$ . Hence,  $(\bar{v} + \bar{e}_i)_i = \bar{v}_i + 1 = t < \bar{w}_i$  when  $i \in \mathcal{L}_1$  and the condition  $\text{Feas}_{i,t+1-1}$  holds for  $(\bar{v} + \bar{e}_i, \bar{w})$ . Since the condition (C1) holds, we have that  $i + 1 \in \mathcal{L}_1$  and  $(\bar{v} + \bar{e}_i)_i = t \leq \bar{v}_{i+1} = (\bar{v} + \bar{e}_i)_{i+1}$ , which implies  $\bar{v} + \bar{e}_i \in \bar{\mathcal{V}}$ . Note that  $(\bar{v} + \bar{e}_i)_\ell = \bar{v}_\ell \geq \bar{v}_{i+1} \geq t = (\bar{v} + \bar{e}_i)_i \forall \ell \in [i + 1, L] \cap \mathcal{L}_1$ . The condition  $\text{Feas}_{i,t+1-2}$  holds for  $(\bar{v} + \bar{e}_i, \bar{w})$ .

The third case holds obviously if  $t = T$ . As shown before, the condition (C2) implies that the amount shipped into node  $(i, t)$  is fully sent to node  $(i + 1, t)$ . We only need to show that  $F_{i+1,t}(\bar{v}, \bar{w})$  is well defined. By the definition,  $(\bar{v}, \bar{w})$  is legitimate from level  $i + 1$  because  $(\bar{v}, \bar{w})$  is legitimate from level  $i$ . We can assume that condition (C1) does not hold, otherwise it becomes the last case, which will be given later. It implies that if  $i + 1 \in \mathcal{L}_1$ , then  $\bar{v}_{i+1} = t - 1$ . Because the condition (C2) holds,  $\bar{w}_{i+1} \geq \bar{w}_i = \bar{v}_i + 1 = t > t - 1 = \bar{v}_{i+1}$  if  $i + 1 \in \mathcal{L}_1$ . The condition  $\text{Feas}_{i+1,t-1}$  holds for  $(\bar{v}, \bar{w})$ . The condition  $\text{Feas}_{i+1,t-2}$  holds obviously for  $(\bar{v}, \bar{w})$  because the condition  $\text{Feas}_{i,t-2}$  holds for  $(\bar{v}, \bar{w})$ .  $\square$

In Proposition 1, the application of the shipment pattern gives dynamic programming recursions on  $F_{i,t}(\bar{v}, \bar{w})$ , which solve the MULS with concave cost functions. Apparently, the bottleneck of the complexity is in recursion (2) where we have  $O(T^{3L_1})$  choices of  $\bar{v}, \bar{u}, \bar{w}$  as they are vectors in  $\bar{\mathcal{V}}$  and  $O(LT)$  choices of indices  $i, t$ . With all these parameters given, the recursion (2) can be evaluated in constant time. Therefore, we have:

**THEOREM 2.** *The MULS with concave cost functions can be solved in  $O(LT^{3L_1+1})$  time.*

**REMARK 1.** In the case of MULS-F, we have  $L_1 = 1$ , so the computation complexity is  $O(LT^4)$ , which is the result derived by Zangwill (1969). More importantly, we have shown that the complexity of MULS depends only on how many echelons have demands, regardless of which echelons have demands.

### 3.2. Fixed-charge cost structure

As one of the most important cost structures, the fixed-charge cost structure receives much attention in literature on studying dynamic lot-sizing problems. For example, Zhang et al. (2012) show that 2-ULS with the fixed-charge cost structure can be solved in  $O(T^4)$  and also give an extended formulation with  $O(T^4)$  variables and  $O(T^4)$  constraints. It is well-known that linear

holding costs can be dropped and replaced by variable production and transportation costs because  $I_t^i = \sum_{\tau=1}^t (x_\tau^i - x_\tau^{i+1}) - d^i(1, t) \forall i \in [1, L-1]$  and  $I_t^L = \sum_{\tau=1}^t x_\tau^L - d^L(1, t)$ . So, in this subsection, we can assume that the holding costs are 0 without loss of generality.

In the case of single-echelon lot-sizing problem with fixed-charge cost structure, the idea of regeneration points and intervals are used to reformulate the problem as a shortest path problem on a directed graph. We will generalize regeneration points and intervals to multiechelon lot-sizing problem as regeneration vectors and arcs, which can be used to reformulate the problem as finding a shortest path on a directed graph. This novel approach is very different from the traditional method that considers MULS as a minimal cost network flow on a two-dimensional grid. We show that the resulting dynamic programming recursions outperform many of the best known algorithms.

**DEFINITION 1.** A vector  $v \in [0, T]^L$  is a regeneration vector if  $v \in \mathcal{V}$  and  $I_{v_i}^i = 0$  for all  $i \in [1, L]$ . Let  $\mathcal{A}_i = \{(v, w) : v, w \in \mathcal{V}, v_j = w_j \forall j \neq i \text{ and } v_i < w_i\}$  for all  $i \in [1, L]$ . A pair of two regeneration vectors  $(v, w)$  forms a regeneration arc at echelon  $i$  with  $i \in [1, L]$ , if  $(v, w) \in \mathcal{A}_i$ .

It is clear that, in the single-echelon case, the regeneration vector becomes a regeneration point and the regeneration arc is a regeneration interval. The definition generalizes an important concept for the multiechelon lot-sizing problem.

Consider a directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ , where the arc set  $\mathcal{A} = \bigcup_{i=1}^L \mathcal{A}_i$ . We define the cost of arc  $(v, w) \in \mathcal{A}_i$  for any  $i \in [1, L]$  as

$$C(v, w) = d^i(v_i + 1, w_i) \sum_{\ell=1}^i c_{v_\ell+1}^\ell + \begin{cases} f_{v_i+1}^i & \text{if } d^i(v_i + 1, w_i) + \sum_{\ell=i+1}^L d^\ell(1, v_\ell) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The arc cost includes the production and transportation costs at all the echelons  $\ell \in [1, i]$  in period  $v_\ell + 1$  for the demand  $d^i(v_i + 1, w_i)$ , and the fixed cost at echelon  $i$  in period  $v_i + 1$  if necessary. The costs can be decomposed into each potential regeneration arc because the variable costs are additive. One of our main theorems shows that:

**THEOREM 3.** *The MULS with fixed-charge cost structure can be solved as a shortest path problem on an acyclic directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{A})$  with the source node  $\mathbf{0}$  and the sink node  $\mathbf{T}$ .*

Before presenting the proof, we highlight two properties. Because MULS with the fixed-charge cost structure is a single source uncapacitated fixed-charge network flow problem, we can apply the well known result that the extreme points of such problem correspond to a spanning tree (see Zangwill (1968), Veinott (1969)) to conclude that

**PROPERTY 1 (Zero Inventory Ordering (ZIO))** *There exists an optimal basic feasible solution with  $I_{t-1}^i x_t^i = 0$  for all  $i \in [1, L]$  and  $t \in [1, T]$ .*

The next property follows from the ZIO property.

**PROPERTY 2** *Suppose that  $I_s^i = I_t^i = 0$  and the inventory quantities between time periods  $s+1$  and  $t$  at echelon  $i$  are all positive. Then,  $x_{s+1}^i = d^i(s+1, t) + \sum_{\tau=s+1}^t x_\tau^{i+1} \forall i \in [1, L-1]$  and  $x_{s+1}^i = d^i(s+1, t)$  when  $i = L$ .*

*Proof* From the inventory balance constraints (1b) and (1c), we get

$$I_t^i = \begin{cases} \sum_{\tau=1}^t (x_\tau^i - x_\tau^{i+1}) - d^i(1, t) & \forall i \in [1, L-1] \\ \sum_{\tau=1}^t x_\tau^i - d^i(1, t) & i = L \end{cases}$$

Thus, for all  $i \in [1, L-1]$ ,

$$\begin{aligned} 0 &= I_s^i - I_t^i = \left( \sum_{\tau=1}^s (x_\tau^i - x_\tau^{i+1}) - d^i(1, s) \right) - \left( \sum_{\tau=1}^t (x_\tau^i - x_\tau^{i+1}) - d^i(1, t) \right) \\ &= x_{s+1}^i - d^i(s+1, t) - \sum_{\tau=s+1}^t x_\tau^{i+1} \end{aligned}$$

Note that Property 1 implies  $x_\tau^i = 0$  for  $\tau = s+2, \dots, t$ . Similarly, we can prove the case when  $i = L$ .  $\square$

Then, we are ready to prove Theorem 3.

*Proof of Theorem 3* Given an extreme solution, the time horizon  $[0, T]$  at each echelon  $i \in [1, L]$  is partitioned by regeneration intervals, which are composed of  $q_i$  regeneration points  $0 = s_{i1} < s_{i2} < \dots < s_{iq_i} = T$  since the starting and ending inventory quantities are 0 by the problem definition. Thus, the total demand can be partitioned by demands on regeneration intervals and Property 2 implies that the demands on each regeneration interval are fulfilled by one production or transportation. For example, for echelon  $i$  and  $k \in [1, q_i]$ , Property 2 implies

$$x_{s_{ik}+1}^i = d^i(s_{ik} + 1, s_{i,k+1}) + \sum_{\tau=s_{ik}+1}^{s_{i,k+1}} x_\tau^{i+1}, \quad (3)$$

i.e., the demands  $d^i(s_{ik} + 1, s_{i,k+1})$  are satisfied by  $x_{s_{ik}+1}^i$ . To realize the demands  $d^i(s_{ik} + 1, s_{i,k+1})$  and the associated cost in a path in  $\mathcal{G}$ , we have two regeneration vectors  $v, w$  such that  $v_i = s_{ik}$  and  $w_i = s_{i,k+1}$ . For other components of  $v, w$ , we have  $v_\ell = w_\ell \forall \ell \in [1, L] \setminus \{i\}$  and they can be obtained iteratively as follows. First, if  $i > 1$ , then we have an index  $j$  such that  $s_{i-1,j} \leq s_{ik} < s_{i-1,j+1}$  and we set  $v_{i-1} = w_{i-1} = s_{i-1,j}$ . The index  $j$  is unique because the regeneration intervals at each echelon give partitions of  $[0, T]$ . Applying Property 2 on echelon  $i-1$  and periods  $s_{i-1,j}, s_{i-1,j+1}$ , we have

$$x_{s_{i-1,j}+1}^{i-1} = d^{i-1}(s_{i-1,j} + 1, s_{i-1,j+1}) + \sum_{\tau=s_{i-1,j}+1}^{s_{i-1,j+1}} x_\tau^i. \quad (4)$$

Because  $s_{i-1,j} \leq s_{ik} < s_{i-1,j+1}$ , the choice of  $v_{i-1}(=s_{i-1,j})$  indicates that the shipment quantity  $x_{v_i+1}^i(=x_{s_{ik}+1}^i)$  is fulfilled by the production or shipment quantity  $x_{v_{i-1}+1}^{i-1}(=x_{s_{i-1,j}+1}^{i-1})$ . Thus, we need to count the shipment cost  $d^i(v_i+1, w_i)c_{v_{i-1}+1}^{i-1}$  for the demand  $d^i(v_i+1, w_i)$ . This procedure can be applied iteratively to all echelons  $1, \dots, i-1$  and the total cost is

$$d^i(v_i+1, w_i) \sum_{\ell=1}^{i-1} c_{v_\ell+1}^\ell + d^i(v_i+1, w_i)c_{v_i+1}^i + f_{v_i}^i$$

where the first term is the aggregated production and transportation costs from all echelons  $1, \dots, i-1$ , the second term is the transportation cost at echelon  $i$  and  $f_{v_i}^i$  is the fixed cost if  $x_{s_{ik}+1}^i$  is positive. Next, if  $i < L$ , then we have a unique index  $r$  (the uniqueness follows a similar reasoning as above) such that  $s_{i+1,r-1} < s_{i,k+1} \leq s_{i+1,r}$  and we set  $v_{i+1} = w_{i+1} = s_{i+1,r}$ . From equation (3) and  $s_{i+1,r-1} < s_{i,k+1} \leq s_{i+1,r}$ , we know the quantity  $x_{s_{i+1,r-1}+1}^{i+1}$  is coming from  $x_{s_{ik}+1}^i$  and  $x_{s_{i+1,r}+1}^{i+1}$  is not. Thus, the choice of  $v_{i+1}$  indicates that the transportation quantities  $x_1^{i+1}, \dots, x_{v_{i+1}}^{i+1}$  and demands  $d^{i+1}(1, s_{i+1,r})$  are fulfilled by the production and transportation quantities  $x_1^i, \dots, x_{v_i}^i$ . Again, iteratively applying the procedure to all the lower echelons of  $i$ , we have regeneration vectors  $v, w$  built and know that the production and transportation quantities  $x_1^i, \dots, x_{v_i}^i$  have satisfied demands  $\sum_{\ell=i+1}^L d^\ell(1, v_\ell)$ , i.e., when we build the path recursively, the demands  $\sum_{\ell=i+1}^L d^\ell(1, v_\ell)$  have already been fulfilled.

Now, we show that the fixed cost defined in the arc cost is valid. Note that if  $x_{v_i+1}^i = x_{s_{ik}+1}^i > 0$ , then  $d^i(v_i+1, w_i) + \sum_{\ell=i+1}^L d^\ell(1, v_\ell) > 0$  obviously. So the fixed cost  $f_{v_i}^i$  will be definitely included into arc cost of  $(v, w)$  if it is necessary. We need to check if the fixed cost defined in the arc cost is valid when  $x_{s_{ik}+1}^i = 0$  but  $d^i(v_i+1, w_i) + \sum_{\ell=i+1}^L d^\ell(1, v_\ell) \neq 0$ . Apparently, in such case, the fixed cost will be considered into the arc cost along the path, even though it is not needed. However, this issue will be resolved easily since we are only interested in the shortest path. If  $x_{s_{ik}+1}^i = 0$ , then we drop  $s_{ik}$  from our selected regeneration points and the quantity of  $x_{s_{i,k-1}+1}^i$  is used to fulfill  $x_{s_{ik}+1}^i$ , which is possible because  $x_{s_{ik}+1}^i = 0$ . Then, following the same procedure discussed above, we derive a similar path with a lower total arc cost with the fixed cost  $f_{v_i}^i(=f_{s_{ik}}^i)$  dropped, since  $s_{ik}$  is dropped.

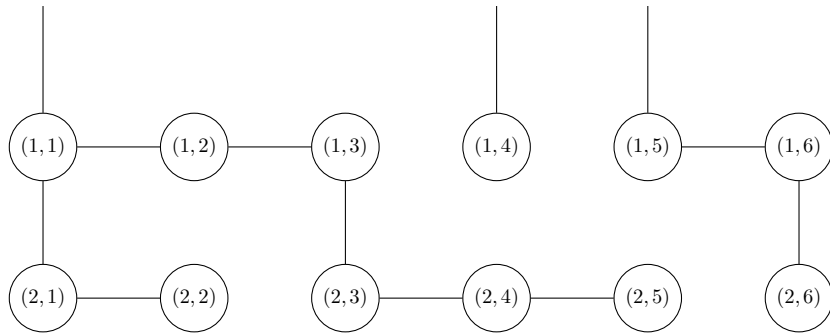
Now, given a path in the graph  $\mathcal{G}$ , simply reversing the procedure described above, we can get a solution from the path that fulfills all the demands and the corresponding objective value has correct production and transportation costs because of the definition of arc costs. As per the previous discussion, the solution may incur unnecessary fixed costs, but this issue is resolved as a result of the minimization of the total cost.

On the one hand, any extreme solution can be converted to a path in  $\mathcal{G}$  with the objective value equal to the total arc costs of the path. On the other hand, any path can be viewed as a solution of

MULS problem and the total arc costs of the path gives the objective value. Therefore, the MULS problem with minimal total cost can be solved as a shortest path problem.  $\square$

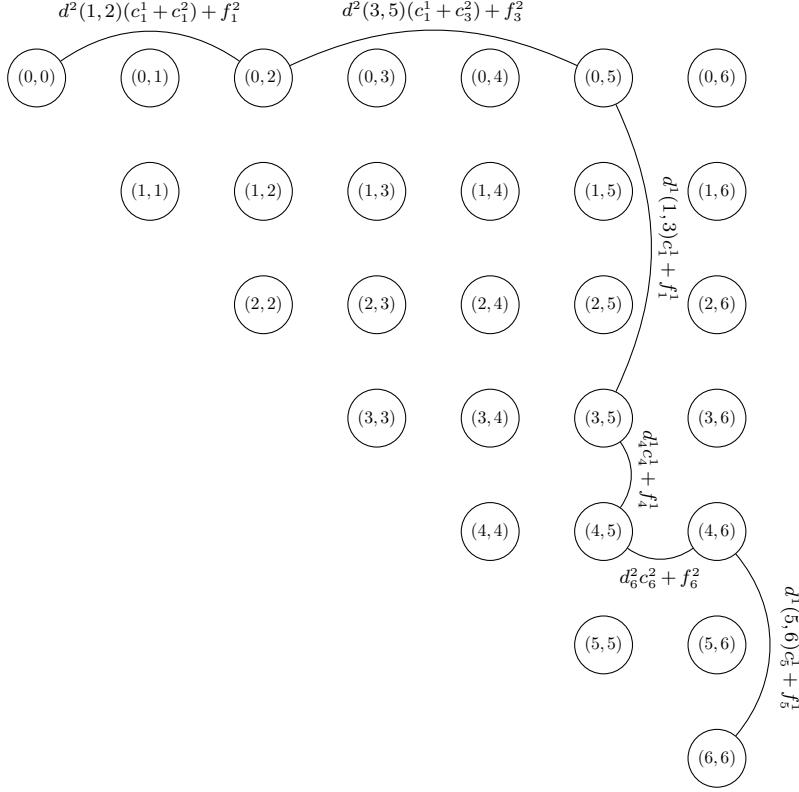
Next we give an example to show that an extreme solution is related to a path composed of regeneration vectors and arcs.

EXAMPLE 1. Given a two-echelon six-period uncapacitated lot-sizing problem, we can view it as a flow network in a two-dimensional grid as in Figure 4. Accordingly, we have a graph  $\mathcal{G}$  in Figure 5 with all vectors in  $\mathcal{V}$ . As shown in Figure 5, the extreme solution in Figure 4 corresponds to a path with cost associated at each arc. To simplify the notation, in this example we assume all demands are positive at each echelon and period. So the fixed cost needs to be considered for each arc in the path.



**Figure 4** An extreme solution of a two-echelon six-period uncapacitated lot-sizing problem.

REMARK 2. In the example, it is clear that the arc  $((0,0), (0,2))$  indicates that demand  $d^2(1,2)$  is satisfied by production and transportation at period 1, the arc  $((0,2), (0,5))$  indicates that demand  $d^2(3,5)$  is satisfied by production at period 1 and transportation at period 3, and the arc  $((0,5), (3,5))$  indicates that demand  $d^1(1,3)$  is satisfied by production at period 1. A question that may arise is what if we have arc  $((0,5), (2,5))$  instead of arc  $((0,5), (3,5))$ . It simply implies that the total production at period 1 is  $d^1(1,2) + d^2(1,5)$  (i.e.,  $d_3^1$  is not in it). Thus, we should have a positive production at period 3 to fulfill the demand  $d_3^1$ . This solution with arc  $((0,5), (2,5))$  indicates that after the production at period 1, we hold amount  $d^2(3,5)$  and transport it to the period 3. From the discussion, we see that, with arc  $((0,5), (2,5))$ , we still have a feasible production and transportation plan. However, the solution is not extreme, because  $I_2^1 x_3^1 > 0$  and the ZIO property is violated. Since the proof shows that the total cost of a path is equal to the total cost of the corresponding solution, the shortest path will correspond to an extreme solution with minimal cost.



**Figure 5** A path in the graph  $\mathcal{G}$  corresponds to the extreme solution in Figure 4.

To formulate MULS with the fixed-charge cost structure, we need to introduce binary variables  $y_t^i$   $\forall i \in [1, L], t \in [1, T]$  to model fixed-charge costs and add the linking constraints  $x_t^i \leq \sum_{\ell=1}^L d^\ell(t, T) y_t^i$ . The shortest path extended formulation for MULS requires variables  $0 \leq \gamma_{v,w} \leq 1$  to indicate the flow on arc  $(v, w) \in \mathcal{A}$ , and can be presented as follows,

$$\begin{aligned}
 \min \quad & \sum_{(v,w) \in \mathcal{A}} C(v,w) \gamma_{(v,w)} \\
 \text{s.t.} \quad & \sum_{v:(v,w) \in \mathcal{A}} \gamma_{(v,w)} - \sum_{v:(w,v) \in \mathcal{A}} \gamma_{(w,v)} = \begin{cases} -1 & \text{if } v = \mathbf{T} \\ 1 & \text{if } v = \mathbf{0} \\ 0 & \text{otherwise} \end{cases} \\
 & 0 \leq \gamma_{(v,w)} \leq 1 \quad \forall (v,w) \in \mathcal{A} \\
 & x_t^i = \sum_{\ell=i}^L \sum_{(v,w) \in \mathcal{A}_\ell: v_i=t} d^\ell(v_\ell + 1, w_\ell) \gamma_{(v,w)} \quad \forall i \in \mathcal{L}, t \in \mathcal{T} \\
 & y_t^i = \sum_{(v,w) \in \mathcal{A}_i: v_i=t} \gamma_{(v,w)} \quad \forall i \in \mathcal{L}, t \in \mathcal{T}.
 \end{aligned}$$

Additionally, considering that there are  $O(T^L)$  nodes and  $O(LT^{L+1})$  arcs in the graph  $\mathcal{G}$ , we have the following corollary.

COROLLARY 1. *The MULS with the fixed-charge cost structure can be solved in  $O(LT^{L+1})$  time, and it has an extended formulation with  $O(LT^{L+1})$  variables and  $O(T^L)$  constraints.*

Note that the arc cost calculation involves  $\sum_{\ell=1}^i c_{v_{\ell+1}}^\ell$  and  $\sum_{\ell=i+1}^L d^\ell(1, v_\ell)$ , which cannot be performed in constant time directly. However, for all  $v \in \mathcal{V}$  and  $i \in [1, L]$ , we can have those two summations calculated in  $O(LT^L)$  time. After this preprocessing, the arc cost can be obtained in constant time and the complexity in Corollary 1 follows the complexity of the shortest path algorithm immediately.

For a given vector  $v \in \mathcal{V}$ , let  $G(v)$  represent the minimum cost of a path in  $\mathcal{G}$  from  $v$  to  $\mathbf{T}$ . For notational convenience, we denote  $v_{L+1} = T$  and  $(v_{-i}, \alpha) \in \mathcal{V}$  as a vector equal to  $v$  except that the  $i$ th component is  $\alpha$ . The shortest path algorithm implies

$$\begin{aligned} G(v) &= \min_{w:(v,w) \in \mathcal{A}} C(v, w) + G(w) \\ &= \min_{i \in [1, L]: v_i < v_{i+1}} \min_{w:(v,w) \in \mathcal{A}_i} C(v, w) + G(w) \end{aligned} \quad (5)$$

$$= \min_{i \in [1, L]: v_i < v_{i+1}} \min_{\substack{w: w_{-i} = v_{-i} \text{ and} \\ v_i < w_i \leq v_{i+1}}} C(v, w) + G(v_{-i}, w_i). \quad (6)$$

Since  $(v, w) \in \mathcal{A}$ , we know  $(v, w) \in \mathcal{A}_i$  for an  $i \in [1, L]$ . From the definition of  $\mathcal{A}_i$  and  $\mathcal{V}$ , we have  $v_{i+1} = w_{i+1} \geq w_i > v_i$ . Thus equations (5) and (6) hold. The arc cost  $C(v, w)$  equals to a possible fixed cost plus  $d^i(v_i + 1, \mu) \sum_{\ell=1}^i c_{v_{\ell+1}}^\ell$ , where the fixed cost  $f_{v_{i+1}}^i$  occurs if we have production or transportation at echelon  $i$  in period  $v_i + 1$ . Apparently, the fixed cost has to be considered if  $d_{v_{i+1}}^i > 0$ . Otherwise, in the case of  $d_{v_{i+1}}^i = 0$ , we can compare the costs by having  $G(v) = G(v_{-i}, v_i + 1)$  (i.e., ignoring the 0 demand) or having production/transportation in period  $v_i + 1$  at echelon  $i$ . Note that  $(v_{-i}, v_i + 1) \in \mathcal{V}$  because  $v_{i+1} > v_i$  in (6). Different from the arc cost, here the way of dealing with fixed cost follows Wagelmans et al. (1992). Therefore,  $G(v)$  satisfies a dynamic programming recursion as follows.

$$G(v) = \min_{i \in [1, L]: v_i < v_{i+1}} \begin{cases} \min_{v_i < \lambda \leq v_{i+1}} \left[ f_{v_{i+1}}^i + d^i(v_i + 1, \lambda) \sum_{\ell=1}^i c_{v_{\ell+1}}^\ell + G(v_{-i}, \lambda) \right] & \text{if } d_{v_{i+1}}^i > 0 \\ \min \left\{ G(v_{-i}, v_i + 1), \min_{v_i < \lambda \leq v_{i+1}} \left[ f_{v_{i+1}}^i + d^i(v_i + 1, \lambda) \sum_{\ell=1}^i c_{v_{\ell+1}}^\ell + G(v_{-i}, \lambda) \right] \right\} & \text{if } d_{v_{i+1}}^i = 0. \end{cases} \quad (7)$$

THEOREM 4. *The MULS with fixed-charge cost structure can be solved in  $O(LT^L \log T)$  time if  $L \leq 3L_1$ , otherwise the complexity follows Theorem 2.*

*Proof* For a given  $i \in [1, L]$  and sub-vector  $v_{-i}$ , we can denote  $G(v_{-i}, \cdot)$  as  $G'(\cdot)$ . Then the recursion (7) becomes the same 1–dimension recursion as the one by Wagelmans et al. (1992)

(equation (1) in their paper), whose value can be determined in  $O(\log T)$  time. Similar results are obtained by Federgruen and Tzur (1991), Aggarwal and Park (1993). Because  $G(v)$  can be obtained in  $O(\log T)$  time with given vector  $v$  and index  $i$ , the overall complexity is  $O(LT^L \log T)$ . If  $L \leq 3L_1$ , this complexity is better than the result in Theorem 2.  $\square$

REMARK 3. By following the traditional network on a two-dimensional grid, the dynamic recursions developed in Melo and Wolsey (2010) and Zhang et al. (2012) consider the regeneration points one echelon after another. Distinct from the traditional approach, our algorithm is to find a shortest path on an  $L$ -dimensional graph with node set  $\mathcal{V}$  such that a node on the shortest path is a regeneration vector which has regeneration points at each echelon as its components. The novel approach allows us to determine the optimal solutions more efficiently. As evidence of the efficacy of our algorithm, Theorem 4 suggests improved complexities than Zangwill (1969)'s for 2-ULS-F and 3-ULS-F, and Zhang et al. (2012)'s for 2-ULS (the complexity comparisons are shown in Table 2). An interesting observation is that the intermediate demands may have no effect on the complexity for 2-ULS, because we have shown the complexity for 2-ULS is  $O(T^2 \log T)$  and the best known complexity for 2-ULS-F is also  $O(T^2 \log T)$  by Melo and Wolsey (2010).

### 3.3. Two-echelon case with backlogging and outbound transportation

The efficiency of our algorithm in Section 3.2 relies on the fact that, without considering the fixed-charge, the production and transportation costs are linear and additive. Hence the total cost can be decomposed into the cost of each arc. Such a property usually does not hold for general cost structures. However, if the cost functions lead to an enumeration of demand units over a set within a reasonable size, then the proposed dynamic programming recursions in Section 3.2 can still be applied to achieve an efficient algorithm.

We consider a 2-ULS with backlogging and outbound transportation, which is studied by Lee et al. (2003) to consolidate outbound transportation scheduling from a third-party warehouse (TPW) to a group of downstream distribution centers (DC). The outbound transportation cost from the TPW consists of a fixed cost and a per-cargo cost. Because of the cargo capacity, the transportation cost follows a stepwise function, which is non-concave. The TPW has the option of releasing an outbound transportation earlier or later than it is actually needed at the DC at the expense of inventory holding or customer waiting cost. The problem can be formulated as follows.

$$\min \sum_{t=1}^n (p_t^1(x_t^1) + p_t^2(x_t^2) + h_t^2 I_t^2 + w_t^2 I_t^-) \quad (8a)$$

$$\text{s.t. } I_{t-1}^1 + x_t^1 = I_t^1 + x_t^2 + d_t^1 \quad \forall t = 1, \dots, T \quad (8b)$$

$$(I_{t-1}^2 - I_{t-1}^-) + x_t^2 = (I_t^2 - I_t^-) + d_t^2 \quad \forall t = 1, \dots, T \quad (8c)$$

$$I_0^i = I_n^i = I_0^- = I_n^- = 0 \quad \forall i = 1, 2 \quad (8e)$$

$$x_t^i \geq 0, I_t^i \geq 0, I_t^- \geq 0 \quad \forall i = 1, 2, t = 1, \dots, T \quad (8f)$$

where the constraint (8c) models the backlogs and variable  $I_t^-$  indicates backlogging quantities in period  $t$ . In the model, for all periods  $t \in [1, T]$ , we have the fixed-charge production costs  $p_t^1(x_t^1) = f_t^1 \delta(x_t^1) + c_t^1 x_t^1$ , the transportation costs  $p_t^2(x_t^2) = f_t^2 \delta(x_t^2) + c_t^2 x_t^2 + A \lceil \frac{x_t^2}{W} \rceil$ , where  $A$  is the cost of delivering a cargo and  $W$  is the capacity of each cargo, inventory holding cost  $h_t^2$  and customer waiting cost  $w_t^2$ . Our model generalizes the one proposed by Lee et al. (2003) since ours includes the variable costs of production and transportation, and intermediate demands. Note that the linear inventory holding costs of  $I_t^1$  at the first echelon are omitted without loss of generality because  $I_t^1 = \sum_{\tau=1}^t (x_\tau^1 - x_\tau^2)$ . Lee et al. (2003) show that an optimal solution can be obtained in  $O(T^6)$  time for the case that demand is occurring at the final echelon only. The main contribution in this subsection is that we are able to obtain an optimal solution in  $O(T^5)$  time even with intermediate demands.

Let  $Q(t, s_1, \beta_1)$ , for all  $t, s_1 \in [1, T]$  and  $0 \leq \beta_1 \leq d_{s_1}^2$ , be the minimal cost of satisfying

- the demands in periods  $t$  through  $T$  at echelon 1 (i.e.,  $d_t^1, \dots, d_T^1$ );
- $d_{s_1}^2 - \beta_1$  unit of demand in period  $s_1$  plus the demands in periods  $s_1 + 1$  through  $T$  at echelon 2 (i.e.,  $d_{s_1+1}^2, \dots, d_T^2$ )

by productions from period  $t$  to  $T$ . We have boundary condition  $Q(T+1, T, d_T^2) = 0$ . The 2-ULS with backlogging and outbound transportation is solved by finding  $Q(1, 1, 0)$ . Note that we could have  $s_1 < t$ , which indicates that the demand in period  $s_1$  at echelon 2 is backlogged. Lee et al. (2003) show that, due to the optimality property, the choices of  $(s_1, \beta_1)$  can be limited into a set, denoted as  $\Gamma$ , with its size of  $O(T^2)$ .

First we have a dynamic recursion when  $(s_1, \beta_1) \neq (T, d_T^2)$ ,

$$Q(t, s_1, \beta_1) = \min_{(s_2, \beta_2) \in \Gamma^*} Q(t, s_2, \beta_2) + g(s_1, \beta_1; s_2, \beta_2) + c_t^1 \cdot (d^2(s_1, s_2 - 1) + \beta_2 - \beta_1) \quad (9)$$

where  $\Gamma^* = \{(s_2, \beta_2) \in \Gamma : s_2 \geq \max(s_1, t) \text{ and } \beta_2 > \beta_1 \text{ if } s_2 = s_1\}$ ,

- in the definition of  $\Gamma^*$ ,  $s_2 \geq t$  is required because the backorders from period  $s_1$  to  $t-1$  at echelon 2 are satisfied by a single transportation (follow the Property 2 and 3 in Lee et al. (2003)). It is obvious that we need  $s_2 \geq s_1$  and  $\beta_2 > \beta_1$  if  $s_2 = s_1$ .
- $g(s_1, \beta_1; s_2, \beta_2)$  is the minimal total cost of satisfying
  - $d_{s_1}^2 - \beta_1$  unit of demand in period  $s_1$  at echelon 2;
  - the demands in periods  $s_1 + 1$  through  $s_2 - 1$  at echelon 2; and
  - $\beta_2$  unit of demand in period  $s_2$  at echelon 2,

i.e., the total demand is  $d_{s_1}^2 - \beta_1 + d^2(s_1 + 1, s_2 - 1) + \beta_2 = d^2(s_1, s_2 - 1) + \beta_2 - \beta_1$ , such that the solution to  $g(s_1, \beta_1; s_2, \beta_2)$  has no regeneration point at the second echelon from period  $s_1$  to  $s_2 - 1$  if  $\beta_1 < d_{s_1}$  and from period  $s_1 + 1$  to  $s_2 - 1$  if  $\beta_1 = d_{s_1}$ .

- the last term of the recursion indicates the variable production cost in period  $t$  for satisfying demands  $d^2(s_1, s_2 - 1) + \beta_2 - \beta_1$  that is transported to the second echelon.

Apparently, the recursion (9) is not enough to obtain the value of  $Q(1, 1, 0)$  since it only works on satisfying demands  $d^2(s_1, s_2 - 1) + \beta_2 - \beta_1$  at echelon 2 from a given production period  $t$ . That is also the reason why the recursion (9) is not necessary when  $(s_1, \beta_1) \neq (T, d_T^2)$ . The next recursion links two production periods  $t$  and  $t'$  is

$$Q(t, s_1, \beta_1) = \min_{t' \in \Delta_{t, s_1, \beta_1}} Q(t', s_1, \beta_1) + \begin{cases} 0 & \text{if } (s_1, \beta_1) = (1, 0) \\ f_t^1 + c_t^1 \cdot d^1(t, t' - 1) & \text{otherwise.} \end{cases} \quad (10)$$

where  $\Delta_{t, s_1, \beta_1} = \{t' \in [t + 1, T + 1] : t \leq T \text{ when } (s_1, \beta_1) \neq (T, d_T^2)\}$ . In the definition of  $\Delta_{t, s_1, \beta_1}$ ,  $t' = T + 1$  only if  $(s_1, \beta_1) = (T, d_T^2)$  and when  $t' = T + 1$  it invokes the boundary condition  $Q(T + 1, T, d_T^2) = 0$ . For example,  $Q(t, T, d_T^2) = Q(T + 1, T, d_T^2) + f_t^1 + c_t^1 \cdot d^1(t, T)$  indicates that period  $t$  is the final period with positive productions and  $Q(t, T, d_T^2)$  is the cost of fulfilling demands from period  $t$  to period  $T$  at echelon 1. The recursion (9) needs to be applied afterward to figure out how many demands at echelon 2 may be satisfied by the production in period  $t$ . The condition  $(s_1, \beta_1) = (1, 0)$  in the recursion (10) indicates that we could have  $G(1, 1, 0) = G(t', 1, 0)$  for some period  $t'$  such that the demands from period 1 to  $t' - 1$  at the second echelon are satisfied by backlogging with production in period  $t'$ . It is clear that we should take the minimal value of those two recursions. In summary, we have

$$Q(t, s_1, \beta_1) = \min \begin{cases} \min_{(s_2, \beta_2) \in \Gamma^*} Q(t, s_2, \beta_2) + g(s_1, \beta_1; s_2, \beta_2) \\ \quad + c_t^1 \cdot (d^2(s_1, s_2 - 1) + \beta_2 - \beta_1) & \text{if } (s_1, \beta_1) \neq (T, d_T^2) \\ \min_{t' \in \Delta_{t, s_1, \beta_1}} Q(t', s_1, \beta_1) + \begin{cases} 0 & \text{if } (s_1, \beta_1) = (1, 0) \\ f_t^1 + c_t^1 \cdot d^1(t, t' - 1) & \text{otherwise.} \end{cases} \end{cases}$$

Suppose that the values of all  $g(s_1, \beta_1; s_2, \beta_2)$  are given. The recursion (9) can be evaluated in constant time with the given  $t, (s_1, \beta_1), (s_2, \beta_2)$ , and (10) can be evaluated in constant time with the given  $t, (s_1, \beta_1), t'$ . Thus, it would be possible to solve the problem in  $O(T^5)$  time. We point out that the above dynamic programming recursions follow the same idea as the recursions (7) in Section 3.2, except that, in the recursion (7), we know the demand in each period is fully satisfied when cost functions are concave, hence the parameter  $\beta_1$  is unnecessary.

Note that  $g(s_1, \beta_1; s_2, \beta_2)$  is similar to the customer subproblem defined by Lee et al. (2003), except that it is independent of any production period. Following a similar approach, we can show that the values of all  $g(s_1, \beta_1; s_2, \beta_2)$  can also be obtained in  $O(T^5)$  time as well, see Appendix B. Therefore, we have

THEOREM 5. *The 2-ULS with backlogging and outbound transportation can be solved in  $O(T^5)$  time.*

#### 4. Capacitated cases

Capacitated lot-sizing problems are usually much more complicated than their uncapacitated counterparts mainly due to the reason that the zero inventory ordering property does not hold after imposing a capacity. However, starting from Florian and Klein (1971), many researchers find that the production and transportation quantities could be enumerated in polynomial time if the capacity is stationary. van Hoesel et al. (2005) provide a detailed study on capacitated multiechelon lot-sizing problem with demand occurring at the final echelon only. They show that, in the view of the network flow problem, the flow corresponding to any extreme solution can be decomposed into a sequence of the so-called subplans, where each subplan has at most one positive production arc strictly less than the capacity. This property of the subplan allows them to enumerate all possible values of cumulative production and transportation quantities in each subplan. Then, they are able to solve MCLS-F through a two-phase dynamic programming by considering adjacent subplans.

A more general approach is proposed in this section to solve MCLS. More specifically, the definition of a subplan needs to be adjusted because of the intermediate demands. It also turns out that, in our case, the cumulative production and transportation quantities have a much richer structure. By enumerating all allowable values, we are able to solve MCLS in  $O(LT^{2L_1L+2})$  time for general concave costs, which outperforms the one proposed by van Hoesel et al. (2005) for solving MCLS-F. Throughout this section, for two  $L$ -dimensional vectors  $u, v$ , we denote  $u \leq v$  if  $u_\ell \leq v_\ell \forall \ell \in [1, L]$  component-wise. Similarly,  $\min(u, v)$  is a  $L$ -dimensional vector whose  $\ell$ th component equals to  $\min(u_\ell, v_\ell) \forall \ell \in [1, L]$ . Recall the definition  $\mathcal{D}_i(\bar{v}, \bar{w}) = \sum_{\ell=i}^L d^\ell(\bar{v}_\ell + 1, \bar{w}_\ell) \forall i \in [1, L]$  in Section 3.1.

##### 4.1. Subplans and relaxed subplans

In this subsection, we will introduce subplans and relaxed subplans for MCLS. We consider flow in the network corresponding to any extreme point, where the network consists of nodes  $(i, t)$  with echelon  $i$  and period  $t$ , see Figure 6. After removing all production arcs, the network is decomposed into several connected components. Then, we connect each isolated node without demand to the node on its left (connecting to its above node if the node is in period 1), and define each connected component as a subplan  $(v, w)$  with  $v, w \in \mathcal{V}$  and  $v \leq w$  if it includes nodes  $(i, v_i + 1), \dots, (i, w_i) \forall i \in [1, L]$  where  $v_{i+1} \leq w_i$  for  $i = 1, \dots, L - 1$ . We also require that, for any  $j \in [1, L - 1]$ , we have either  $v_{j+1} < w_j$  or  $v_\ell = w_\ell \forall \ell \in [j + 1, L]$ , because of the arborescent structure of an extreme solution in the subplan. Note that  $v_{j+1} \geq w_j$  indicates that, in the subplan, there is no flow that

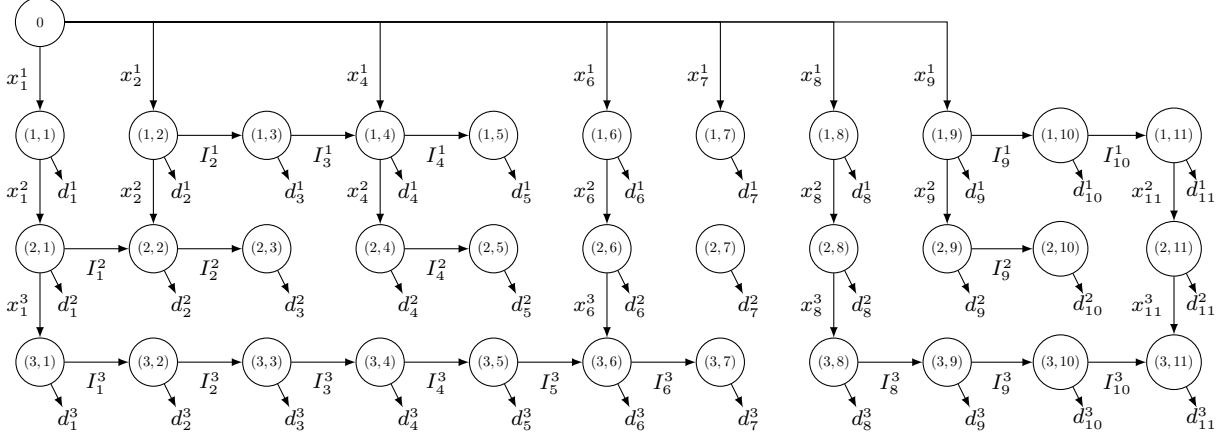
can connect to the nodes  $(j+1, v_{j+1}+1), \dots, (j+1, w_{j+1})$ . Hence, the subplan only contains nodes at echelons  $1, \dots, j$  and we need to have  $v_\ell = w_\ell \forall \ell \in [j+1, L]$ . We say two subplans  $(v', w')$  and  $(v'', w'')$  are consecutive if  $w' = v''$ . For example, in Figure 6, if  $d_7^1 > 0$ , then we have three consecutive subplans  $((0, 0, 0), (6, 7, 7)), ((6, 7, 7), (7, 7, 7))$  and  $((7, 7, 7), (11, 11, 11))$ , otherwise we only have two consecutive subplans  $((0, 0, 0), (7, 7, 7))$  and  $((7, 7, 7), (11, 11, 11))$ . As is common in network flow problems, we define an arc as free if it carries an amount of flow that is both positive and strictly less than its capacity. Because the flow corresponding to the extreme solution is acyclic, there is at most one free production arc entering the subplan.

The definition of the subplan is introduced by van Hoesel et al. (2005) except that their additional requirement is  $v_{i+1} < w_i \forall i \in [1, L-1]$  for a subplan  $(v, w)$ . For example, in Figure 6,  $((6, 7, 7), (7, 7, 7))$  is not a valid subplan anymore because  $v_2 = w_1 = 7$ . Thus, we have only two consecutive subplans  $((0, 0, 0), (7, 7, 7))$  and  $((7, 7, 7), (11, 11, 11))$  when  $d_7^1 > 0$ . Apparently, we could have two free production arcs, for example  $x_4^1$  and  $x_7^1$ , in the subplan  $((0, 0, 0), (7, 7, 7))$ . The example shows that, with the requirements  $v_{i+1} < w_i \forall i \in [1, L-1]$  in the case of having intermediate demands, we cannot assure that there is at most one free production arc entering each subplan, which is the key property we need for subplans. Therefore, we use a different (relaxed) requirement in our subplan definition, which allows a subplan to contain only nodes at some higher levels, such as the subplan  $((6, 7, 7), (7, 7, 7))$ .

In summary, any extreme solution can be decomposed into a sequence of consecutive subplans and there is at most one free production arc entering the subplan. Even though we need two vectors  $v, w \in \mathcal{V}$  for a subplan, the total demand of a subplan depends only on  $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$  where  $\bar{v}_\ell = v_\ell$  and  $\bar{w}_\ell = w_\ell \forall \ell \in \mathcal{L}_1$ . So we define the pair  $(\bar{v}, \bar{w})$  as a relaxed subplan. We know that the subplan  $(v, w)$  satisfies, for any  $j \in [1, L-1]$ , either  $v_{j+1} < w_j$  or  $v_\ell = w_\ell \forall \ell \in [j+1, L]$ . It implies that  $(\bar{v}, \bar{w})$  satisfies, for any  $j \in [1, L-1] \cap \mathcal{L}_1$ , either  $v_{j+1} < w_j$  or  $v_\ell = w_\ell \forall \ell \in [j+1, L]$ . Thus,  $(\bar{v}, \bar{w})$  is a legitimate pair from level 1 defined in Section 3.1. From now on, a relaxed subplan  $(\bar{v}, \bar{w})$  always indicates that  $(\bar{v}, \bar{w})$  is a legitimate pair from level 1. Similarly, two relaxed subplans  $(\bar{v}', \bar{w}')$  and  $(\bar{v}'', \bar{w}'')$  are consecutive if  $\bar{w}' = \bar{v}''$ . Based on the above discussions, we have

**PROPOSITION 2.** *The total demand  $\mathcal{D}_1(\mathbf{0}, \bar{\mathbf{T}})$  can be decomposed into demands of a sequence of consecutive relaxed subplans. The demand of a relaxed subplan is served by at most one positive production quantity whose value is strictly less than the capacity.*

Although the concepts of subplans and relaxed subplans are related, they are fundamentally different. Subplan is a subnetwork connected by a flow that indicates a production and transportation plan of a subproblem, whereas relaxed subplan is just a pair of vectors to indicate the demand quantities. When demand occurs at the final echelon only, a relaxed subplan is identified by a pair  $(\bar{v}, \bar{w})$  whose only non-fixed components are at the final echelon. Then, it is equivalent to represent the relaxed subplan as  $(\bar{v}_L, \bar{w}_L)$ .



**Figure 6** The flow corresponding to an extreme solution to the 3-CLS with 11 periods.

## 4.2. Production and transportation quantities in a subplan

In this subsection, we fix a subplan  $(v, w)$  and the corresponding relaxed subplan  $(\bar{v}, \bar{w})$ . Note that a subplan is a connected network. For an arbitrary time period  $t \in [\bar{v}_1 + 1, T]$ , let  $X_1$  be the cumulative production quantity up to and including period  $t$  in the subplan, and  $X_i, \forall i \in [2, L]$ , be the cumulative transportation quantity from echelon  $i - 1$  to echelon  $i$  up to and including period  $t$  in the subplan. Clearly,  $X_1 \geq \dots \geq X_L \geq 0$ . We define vector  $X = (X_1, \dots, X_L)$  as a cumulative quantity vector to simplify the notation. The purpose of this subsection is to identify all allowable values of  $X_i \forall i \in [1, L]$ .

Let  $K = \lfloor \mathcal{D}_1(\bar{v}, \bar{w}) / \mathbb{C} \rfloor$  and  $\epsilon = \mathcal{D}_1(\bar{v}, \bar{w}) - K\mathbb{C}$ . Apparently,  $0 \leq \epsilon < \mathbb{C}$ . Because of Proposition 2, the subplan  $(v, w)$  has exactly one production flow with  $\epsilon$  amount if  $\epsilon > 0$ . Since  $X_1$  is the cumulative production quantity in the subplan, we have

$$X_1 \in \bigcup_{k=0}^K \{k\mathbb{C}, k\mathbb{C} + \epsilon\} \quad (11)$$

where, in addition,  $\mathcal{D}_1(\bar{v}, \min(t \cdot \bar{1}, \bar{w})) \leq X_1 \leq \min(K\mathbb{C} + \epsilon, (t - \bar{v}_1)\mathbb{C})$  to ensure that (i)  $X_1$  is enough to cover the demands up to and including period  $t$  in the subplan; (ii)  $X_1$  is bounded by the total capacities in periods from  $\bar{v}_1 + 1$  to  $t$  and the total demands in the subplan. Next, we introduce Proposition 3, which is proved in Appendix C, to characterize the possible values for the transportation quantities  $X_i$  for a given  $i \in [2, L]$ .

**PROPOSITION 3.** *The cumulative transportation quantities  $X_i$  with  $i \in [2, L]$  have three allowable forms as follows.*

Form (1): *there exist  $j_1 \in [i - 1, L - 1]$ ,  $\mu_0^1 \in [0, K]$ ,  $\mu_\ell^1 \in [v_\ell + 1, w_\ell] \forall \ell \in [1, j_1]$  and  $\eta_\ell^1 \in [v_\ell + 1, w_\ell] \forall \ell \in [i, j_1]$  such that*

$$X_i = Y(\mu_0^1) - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \mu_\ell^1) - \sum_{\ell=i}^{j_1} d^\ell(\eta_\ell^1, \mu_\ell^1).$$

Form (2): *there exist  $j_2 \in [i-1, L-1]$ ,  $\mu_0^2 \in [0, K]$ ,  $\mu_\ell^2 \in [v_\ell + 1, w_\ell] \forall \ell \in [1, j_2]$  and  $\eta_\ell^2 \in [v_\ell + 1, w_\ell] \forall \ell \in [i, j_2]$  such that*

$$X_i = Y(\mu_0^2) - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \mu_\ell^2) + \sum_{\ell=i}^{j_2} d^\ell(\eta_\ell^2, \mu_\ell^2).$$

Form (3): *there exist  $\mu_\ell^3 \in [v_\ell + 1, w_\ell] \forall \ell \in [i, L]$  such that*

$$X_i = \sum_{\ell=i}^L d^\ell(v_\ell + 1, \mu_\ell^3),$$

where  $Y(\tau) \in \{\tau\mathbb{C}, \tau\mathbb{C} + \epsilon\}$ .

We make a few comments on the proposition,

- (i) When demand occurs only at the final echelon, Forms (1) and (2) degenerate to  $X_i = Y(\tau)$  for some  $\tau \in [0, K]$ . Together with Form (3), Proposition 3 is equivalent to the result shown by van Hoesel et al. (2005). Therefore, Proposition 3 generalizes the results for MCLS-F.
- (ii) The most important property of  $X_i$  shown by Proposition 3 is that, in Forms (1) and (2), besides the total production quantity of an initial sequence of production periods minus the demands of an initial sequence of periods at each above level, i.e.,

$$Y(\mu_0^1) - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \mu_\ell^1) \quad \text{or} \quad Y(\mu_0^2) - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \mu_\ell^2),$$

the value of  $X_i$  can be characterized by the demands in some periods at echelon  $i, \dots, j_1(j_2)$  where (a)  $j_1(j_2) < L$ ; (b) those periods in each echelon are consecutive.

- (iii) We give two examples on Forms (1) and (2). For example, in subplan  $((0, 0, 0), (6, 7, 7))$  of Figure 6, the cumulative transportation quantity up to period 2 at echelon 2,  $x_1^2 + x_2^2$ , equals to the production up to period 5, i.e.,  $x_1^1 + x_2^1 + x_4^1$ , minus an initial sequence of demand periods  $d^1(1, 5)$ , and further minus demand  $d^2(4, 5)$ , as described in Form (1).

Another example is in subplan  $((7, 7, 7), (11, 11, 11))$ , the cumulative transportation quantity up to period 9 at echelon 2,  $x_8^2 + x_9^2$ , is equal to the production up to period 8 in the subplan, i.e.,  $x_8^1$ , minus an initial sequence of demand periods  $d^1(8, 8)$ , and further plus demand  $d^2(9, 10)$ , as described in Form (2).

In addition, we require  $\mathcal{D}_i(\bar{v}, \min(t \cdot \bar{1}, \bar{w})) \leq X_i \leq X_{i-1} - d^{i-1}(\bar{v}_{i-1} + 1, \min(t, \bar{w}_{i-1})) \forall i \in [2, L]$ , and  $X_i = 0$  when  $t \leq \bar{v}_i \forall i \in [2, L] \cap \mathcal{L}_1$  to ensure that (i) the demands are satisfied on time, (ii) the quantity  $X_{i-1}$  is used to first satisfy demands at echelon  $i-1$  up to and including period  $t$  in the subplan, then the part of it is transported to the next echelon  $i$ , and (iii) the transportation is in the relaxed subplan at echelons with intermediate demands. Because there is no demand at any echelon in  $\mathcal{L}_0$ , only the pair  $(\bar{v}, \bar{w})$  representing a relaxed subplan matters in defining  $X_i \forall i \in [1, L]$ . Therefore, all allowable values of  $X$  depend on  $t, \bar{v}$  and  $\bar{w}$ , and we introduce the following definition.

DEFINITION 2. Let  $\Theta_{t,(\bar{v},\bar{w})}$  be the set of all allowable values of  $X$  in the subplan  $(v, w)$ .

Now we prove Proposition 3.

REMARK 4. We can summarize the complexity of allowable values for  $X \in \Theta_{t,(\bar{v},\bar{w})}$  as follows.

- The number of allowable values for  $X_1$  is in  $O(T)$  because it depends on  $K$  and  $K \leq T$ .
- The complexity of the number of allowable values for  $X_i \forall i \in [2, L]$  depends on the number of parameters and the size of the intervals  $[v_\ell + 1, w_\ell] \forall \ell \in \mathcal{L}_1$  in Proposition 3. The number of parameters needed for Form (3) is  $|\mathcal{L}_1 \cap [i, L]|$ . The number of parameters needed for Form (1) is  $1 + |\mathcal{L}_1 \cap [1, i - 1]| + 2|\mathcal{L}_1 \cap [i, j_1]| \leq L_1 + |\mathcal{L}_1 \cap [i, L - 1]|$  because  $j_1 < L$ . We have the same result for Form (2). So, the number of allowable values for  $X_i$  for each  $i \in [2, L]$  is in  $O(T^{L_1 + |\mathcal{L}_1 \cap [i, L - 1]|})$  or  $O(T^{2L_1})$  if relaxing further. From the discussion above, it is easy to calculate that the number of allowable values for  $X$  is  $O(T^{1 + 2L_1(L - 1)})$ .

### 4.3. General concave costs

Given a relaxed subplan  $(\bar{v}, \bar{w})$ , we define  $\phi_{(\bar{v}, \bar{w})}(t, X)$ , with a time period  $t \in [\bar{v}_1, T]$  and cumulative quantity vector  $X \in \Theta_{t,(\bar{v}, \bar{w})}$ , as the the optimal cost where the demand  $\mathcal{D}_1(\mathbf{0}, \bar{v})$  is satisfied, plus the optimal cost of having  $X_1$  as the cumulative production quantity up to and including period  $t$  *in the subplan*, and  $X_i, \forall i \in [2, L]$ , as the cumulative transportation quantity from echelon  $i - 1$  to echelon  $i$  up to and including period  $t$  *in the subplan*. We also require that  $t \cdot \mathbb{C} \geq \mathcal{D}_1(\mathbf{0}, \bar{v})$  because the demand  $\mathcal{D}_1(\mathbf{0}, \bar{v})$  has been satisfied. The goal is to determine the value of  $\phi_{(\bar{\mathbf{T}}, \bar{\mathbf{T}})}(T, \mathbf{0})$ .

Note that  $\phi_{(\bar{v}, \bar{w})}(t, \mathbf{0})$  is simply the optimal cost of satisfying the demand  $\mathcal{D}_1(\mathbf{0}, \bar{v})$ , and  $\phi_{(\bar{v}, \bar{w})}(t, \mathcal{D}_1(\bar{v}, \bar{w}) \cdot \mathbf{1})$  is the optimal cost of satisfying the demand  $\mathcal{D}_1(\mathbf{0}, \bar{w})$ . So, we have

$$\phi_{(\bar{v}, \bar{w})}(t, \mathbf{0}) = \phi_{(\bar{u}, \bar{v})}(t, \mathcal{D}_1(\bar{u}, \bar{v}) \cdot \mathbf{1}), \quad (12)$$

for any relaxed subplan  $(\bar{u}, \bar{v})$  with  $\bar{u} \leq \bar{v}$ . Note that (12) implies that  $t \cdot \mathbb{C} \geq \mathcal{D}_1(\mathbf{0}, \bar{v})$ . We also give boundary conditions  $\phi_{(\mathbf{0}, \mathbf{0})}(\mathbf{0}, \mathbf{0}) = 0$ . With this boundary conditions and (12), we have  $\phi_{(\mathbf{0}, \bar{w})}(\mathbf{0}, \mathbf{0}) = 0 \forall \bar{w} \in \bar{\mathcal{V}}$ .

Suppose that we have  $\phi_{(\bar{v}, \bar{w})}(t - 1, \bar{X})$  with  $\bar{X} \in \Theta_{t-1,(\bar{v}, \bar{w})}$ , i.e.,  $\bar{X}$  is a cumulative quantity vector up to and including period  $t - 1$  in the same relaxed subplan. Then, at echelon  $i \in [1, L]$  in period  $t$ , the production quantity is  $X_i - \bar{X}_i$  and the inventory quantity is  $X_i - X_{i+1} - d^i(\bar{v}_i + 1, \min(t, \bar{w}_i))$ , except that, at echelon  $L$ , the inventory quantity is  $X_L - d^L(\bar{v}_L + 1, \min(t, \bar{w}_L))$ . Apparently,  $X_1 - \bar{X}_1 \in \{0, \epsilon, \mathbb{C}\}$  and  $X_i \geq \bar{X}_i \forall i \in [2, L]$ . So, we have recursion

$$\begin{aligned} \phi_{(\bar{v}, \bar{w})}(t, X) = & \min_{\bar{X} \in \Theta_{t-1,(\bar{v}, \bar{w})}(X)} \phi_{(\bar{v}, \bar{w})}(t - 1, \bar{X}) + \sum_{i=1}^L p_t^i(X_i - \bar{X}_i) \\ & + \sum_{i=1}^{L-1} h_t^i(X_i - X_{i+1} - d^i(\bar{v}_i + 1, \min(t, \bar{w}_i))) + h_t^L(X_L - d^L(\bar{v}_L + 1, \min(t, \bar{w}_L))), \quad (13) \end{aligned}$$

where  $\Theta_{t-1,(\bar{v},\bar{w})}(X) = \{X' \in \Theta_{t-1,(\bar{v},\bar{w})} : X_1 - X'_1 \in \{0, \epsilon, \mathbf{C}\}, X_i \geq X'_i \forall i \in [2, L]\}$ . The MCLS can be solved by recursion (12), (13) and boundary conditions  $\phi_{(\mathbf{0},\mathbf{0})}(0, \mathbf{0}) = 0$ .

Recall that Remark 4 indicates that the number of allowable values for  $X$  is in  $O(T^{1+2L_1(L-1)})$ . With the given  $\bar{v}, \bar{w}, t$ , the recursion (12) can be evaluated in  $O(T^{L_1})$  time because of  $\bar{u}$ , and  $d^i(\bar{v}_i + 1, \min(t, \bar{w}_i)) \forall i \in [1, L]$  can be evaluated in  $O(L)$ . Thus, the total time that recursion (12) needs is in  $O(T^{3L_1+1})$ . If  $X$  and  $\bar{X}$  are given as well, then recursion (13) can be evaluated in  $O(L)$  time. With observation that  $X - \bar{X} \in \{0, \epsilon, \mathbf{C}\}$  and simple calculations, we can evaluate the recursion (13) in  $O(LT^{4L_1L+2L_1+3})$  time. Thus, MCLS can be solved in polynomial time when  $L$  is fixed. The relations between  $X$  and  $\bar{X}$  can be explored further as they are cumulative quantity vectors up to two adjacent time periods in the same relaxed subplan. This can lead to a more efficient evaluation of the recursion (13) with a complexity of  $O(LT^{2L_1L+2})$ . Let  $\Theta_{t,(\bar{v},\bar{w})}^2$  be the set including all allowable values of pairs  $(X, \bar{X})$ . We have the following proposition proved in Appendix D.

PROPOSITION 4. *The number of allowable values in  $\Theta_{t,(\bar{v},\bar{w})}^2$  is in  $O(T^{2L_1(L-1)+1})$ .*

The recursion (13) can be evaluated in  $O(L)$  time with given  $\bar{v}, \bar{w}, t$ , and  $(X, \bar{X}) \in \Theta_{t,(\bar{v},\bar{w})}^2$ . Thus, the complexity is  $O(LT^{2L_1+1+2L_1(L-1)+1}) = O(LT^{2L_1L+2})$ . We define  $\Theta_{t,(\bar{v},\bar{w})}^2(X)$  as the collection of all the allowable values of  $\bar{X}$  with  $(X, \bar{X}) \in \Theta_{t,(\bar{v},\bar{w})}^2$ . Summarizing all the discussions above, we have the following theorem.

THEOREM 6. *The MCLS can be solved by dynamic programming recursions as follows,*

$$\begin{aligned} \phi_{(\bar{v},\bar{w})}(t, X) = & \min_{\bar{X} \in \Theta_{t,(\bar{v},\bar{w})}^2(X)} \phi_{((\bar{v},\bar{w}))}(t-1, \bar{X}) + \sum_{i=1}^L p_t^i(X_i - \bar{X}_i) \\ & + \sum_{i=1}^{L-1} h_t^i(X_i - X_{i+1} - d^i(\bar{v}_i + 1, \min(t, \bar{w}_i))) + h_t^L(X_L - d^L(\bar{v}_L + 1, \min(t, \bar{w}_L))) \end{aligned} \quad (14)$$

with boundary conditions  $\phi_{(\bar{v},\bar{w})}(t, \mathbf{0}) = \phi_{(\bar{u},\bar{v})}(t, \mathcal{D}_1(\bar{u}, \bar{v}) \cdot \mathbf{1}) \forall \bar{u} \leq \bar{v}$  and relaxed subplan  $(\bar{u}, \bar{v})$ , and  $\phi_{(\mathbf{0},\mathbf{0})}(0, \mathbf{0}) = 0$ , where the recursions (14) can be performed in  $O(LT^{2L_1L+2})$  time.

REMARK 5. For 1-CLS, we have  $L = L_1 = 1$ . Theorem 6 implies the same complexity  $O(T^4)$  as Florian and Klein (1971). In the case of 2-CLS-F, i.e.,  $L = 2$  and  $L_1 = 1$ , Theorem 6 indicates that the model can be solved in  $O(T^6)$  time, which improves the complexity  $O(T^7)$  proposed by van Hoesel et al. (2005) and the complexity  $O(T^8)$  proposed by Hwang et al. (2013).

#### 4.4. Fixed-charge transportation costs without speculative motives

The assumption of no speculative motives is commonly assumed for the production and inventory holding costs in traditional economic lot-sizing models and appears often in the literature. In the

context of fixed-charge transportation costs, no speculative motives indicates that it is attractive to transport as late as possible. More formally,  $c_t^\ell + h_t^{\ell+1} \geq c_{t+1}^\ell + h_t^\ell$  if inventory holding cost is linear. With the same arguments by van Hoesel et al. (2005), there exists an optimal solution satisfying the ZIO (zero inventory ordering) property, i.e.,  $I_{t-1}^i x_t^i = 0$  for  $i \in [1, L]$  and  $t \in [1, T]$ , because of fixed-charge transportation costs and no speculative motives. Therefore, we can limit ourselves on the extreme solutions with the ZIO property. The MCLS can be solved by decoupling the production echelon (the first echelon) from the rest of the model and the echelon from level 2 to level  $L$  can be solved by using dynamic programming recursions in Section 3.1 due to the ZIO property. Note that the production costs and inventory holding costs are still assumed to be generally concave.

Given a relaxed subplan  $(\bar{v}, \bar{w})$ , slightly different from Section 4.3, we define  $\psi_{(\bar{v}, \bar{w})}(t, X_1, X_2)$ , with a time period  $t \in [\bar{v}_1, T]$ , as the the optimal cost where the demand  $\mathcal{D}_1(\mathbf{0}, \bar{v})$  is satisfied, plus the optimal cost of having  $X_1$  as the cumulative production quantity up to and including period  $t$  in the subplan, and  $X_2$  as the cumulative transportation quantity from echelon 1 to echelon 2 up to and including period  $t$  in the subplan. The goal is to determine the value of  $\psi_{(\bar{v}, \bar{w})}(T, 0, 0)$ . Because of the ZIO property and the proof of Proposition 1, we have  $X_2 = \mathcal{D}_2(\bar{v}, \bar{\mu})$  for some  $\bar{\mu} \in \Omega_{2,t}(\bar{v}, \bar{w})$  (see the definition in Section 3.1). Without loss of generality, we can redefine the notation as  $\psi_{(\bar{v}, \bar{w})}(t, X_1, \bar{\mu})$ . Following the dynamic programming recursions (14), we have

$$\begin{aligned} \psi_{(\bar{v}, \bar{w})}(t, X_1, \bar{\mu}) = & \min_{\substack{\bar{X}_1 - X_1 \in \{0, \epsilon, \mathbb{C}\} \\ \text{and } \bar{\eta} \in \Omega_{2,t}(\bar{v}, \bar{\mu})}} \psi_{(\bar{v}, \bar{w})}(t-1, \bar{X}_1, \bar{\eta}) + p_t^1(X_1 - \bar{X}_1) \\ & + h_t^1(X_1 - \mathcal{D}_2(\bar{v}, \bar{\mu}) - d^1(\bar{v}_1 + 1, \min(t, \bar{w}_1))) + F_{2,t}(\bar{\eta}, \bar{\mu}) \end{aligned} \quad (15)$$

where  $F_{2,t}(\bar{\eta}, \bar{\mu})$  and  $\Omega_{2,t}(\bar{v}, \bar{\mu})$  are defined in Section 3.1. We have boundary conditions  $\phi_{(\bar{v}, \bar{w})}(t, \mathbf{0}) = \phi_{(\bar{u}, \bar{v})}(t, \mathcal{D}_1(\bar{u}, \bar{v}) \cdot \mathbf{1}) \forall \bar{u} \leq \bar{v}$  and relaxed subplan  $(\bar{u}, \bar{v})$ , and  $\phi_{(\mathbf{0}, \mathbf{0})}(0, \mathbf{0}) = 0$ .

Note that  $F_{2,t}(\bar{\eta}, \bar{\mu})$ , for all  $t, \bar{\eta}, \bar{\mu}$ , can be determined in  $O(LT^{3L_1+1})$  as shown in Section 3.1. Given  $\bar{v}, \bar{w}, t, X_1, \bar{\mu}, \bar{\eta}$  and the value of  $F_{2,t}(\bar{\eta}, \bar{\mu})$ , the recursion can be evaluated in constant time. The complexity is  $O(T^{4L_1+2} + LT^{3L_1+1})$ . Therefore, we have

**THEOREM 7.** *The MCLS with fixed-charge transportation costs and no speculative motives can be solved in  $O(T^{4L_1+2} + LT^{3L_1+1})$ .*

**REMARK 6.** Theorem 7 is different from Theorem 6 in that it shows a complexity which is independent of the number of echelon  $L$ . Under the condition of fixed-charge transportation costs with no speculative motives, the MCLS can be solved within polynomial time  $O(T^{4L_1+2} + LT^{3L_1+1})$  when  $L_1$  is fixed. It generalizes and outperforms the complexity  $O(T^7 + LT^4)$  for MCLS-F, i.e.  $L_1 = 1$ , by van Hoesel et al. (2005). In the case of 2-CLS-F,  $\bar{v}, \bar{w}, \bar{\mu}, \bar{\eta}$  are equivalent to scalars because only

their second components are not fixed to 0 and  $\bar{\mu}_2 \geq t$ . As shown by van Hoesel et al. (2005), we must have  $\bar{X}_2 = X_2$  when  $\bar{\mu}_2 > t$  because no speculative motives implies that we should transport as late as possible. So  $\bar{\eta} = \bar{\mu}$  when  $\bar{\mu}_2 > t$ . This will reduce the complexity to  $O(T^5)$ .

## 5. Conclusions

We study the multiechelon lot sizing with intermediate demands (MLS). Many existing studies have provided polynomial algorithms on MLS with demand occurring at the final echelon only, or shown that the multiple sources network flow problem is NP-hard. However, the complexity of MLS with the fixed-charge cost structure, which is a classical single source network flow problem, remains unknown. As one of many contributions in this paper, we prove that the MLS with the fixed-charge cost structure is NP-hard and close the theoretical gap.

We investigate both uncapacitated and capacitated MLS with different types of cost functions, such as the general concave costs, the fixed-charge cost structure, stepwise and non-concave transportation cost, and fixed-charge transportation cost with no speculative motives. By considering intermediate demands, our results (see Table 1) generalize much of the existing research, such as Zangwill (1969), Lee et al. (2003) and van Hoesel et al. (2005), which are special cases of the problems studied in this paper. In addition to developing efficient algorithms for solving both MULS and MCLS, we show that the complexities of our algorithms outperform that of many best known algorithms in the literature (see Table 2).

## Proofs of Statements

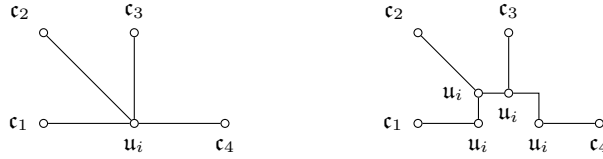
### Appendix A: Proof of NP-hardness

According to the **Overall strategy** section in Section 2, the proof includes three steps in three subsections as follows.

#### A.1. Embedding

Given the identification graph  $\mathfrak{G}$  of an instance of planar 3SAT, we follow the procedure used by Shi and Su (2015) to embed it into a rectilinear grid and denote the embedded graph as  $\mathfrak{R}$ . We first convert the graph  $\mathfrak{G}$  to a planar graph  $\mathfrak{H}$  with maximum degree 3 in two steps:

1. We replace a variable  $u$  by a path of  $\deg(u)$  many duplicated variable vertices, where  $\deg(u)$  is the degree of  $u$ . See Figure EC.1 for an example.



**Figure EC.1** Vertex  $u_i$  of degree 4 is replaced by a path of 4 vertices.

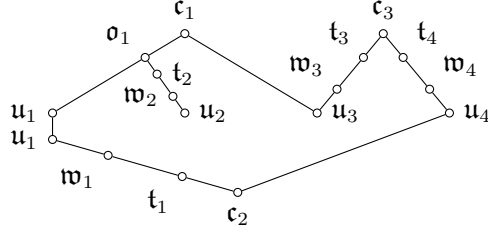
2. Next, we will introduce NOT vertices and OR vertices to facilitate our network design.

For each clause  $c_j \in \mathfrak{G}$  that contains two variables, we maintain the form of  $c_j = u_i \vee \bar{u}_k$ . If  $c_j$  is not in the right form, for example,  $c_j = u_i \vee u_k$ , then we insert a NOT vertex  $\bar{u}_k$  and an auxiliary vertex  $t_k$  between the variable vertex  $u_k$  and the clause vertex  $c_j$ , where  $\bar{u}_k$  denotes the negation operator and  $t_k$ , which is next to the clause vertex, indicates the negation value of the literal  $u_k$ . So, in this case,  $t_k = \bar{u}_k$  and  $c_j = u_i \vee \bar{t}_k$ , which is in the required form.

For each clause  $c_j \in \mathfrak{G}$  that contains three variables, we maintain the form of  $c_j = u_i \vee \bar{u}_k \vee \bar{u}_l = (u_i \vee \bar{u}_k) \vee \bar{u}_l = \sigma_j \vee \bar{u}_l$ . We add an OR vertex according to the newly introduced clause  $\sigma_j$ . Then, we connect  $c_j$  to  $\sigma_j$ , and replace the two edges  $(c_j, u_i)$  and  $(c_j, u_k)$  by  $(\sigma_j, u_i)$  and  $(\sigma_j, u_k)$ . Similarly, if  $c_j$  is not in the desired form, we will insert NOT vertices as needed.

After performing the two-step procedure described above, the graph  $\mathfrak{G}$  for the example in Figure 2 is converted to a graph as shown in Figure EC.2:

As shown by Shi and Su (2015), any planar graph with node set  $\mathfrak{V}$  of maximum degree 3 can be embedded into a rectilinear grid of area  $O(|\mathfrak{V}|^2)$ , and the following properties as shown in Figure EC.3 could be obtained by locally rearranging the grid embedding, increasing the size of grid by a constant factor, adding additional variable vertices with their connecting edges. For each clause



**Figure EC.2** The graph  $\mathfrak{H}$  for the example in Figure 2.

vertex, there are only two possible layouts in the embedded graph  $\mathfrak{R}$  as shown in Figure EC.3, where  $l_{ik}$  (or  $l_{jk}$ ) indicates a variable vertex or auxiliary vertex for  $k = 1, 2, 3$ , and an edge with cross sign indicates that such edge cannot exist in the embedded graph  $\mathfrak{R}$ . We have

- for each clause vertex  $c_i = l_{i1} \vee \bar{l}_{i3}$  in the graph  $\mathfrak{R}$ , the path from  $l_{i1}$  enters from the above, and the path from  $l_{i3}$  enters from the left. See Layout I in Figure EC.3.
- for each clause vertex  $c_j = o_j \vee \bar{l}_{j3} = (l_{j1} \vee \bar{l}_{j2}) \vee \bar{l}_{j3}$  in the graph  $\mathfrak{R}$ , the path from  $o_j$  enters from the above, and the path from  $l_{j3}$  enters from the left. For the OR vertex  $o_j$ , the path from  $l_{j1}$  enters from the above, and the path from  $l_{j2}$  enters from the left. See Layout II in Figure EC.3.



(a) Layout I:  $c_j = l_{j1} \vee \bar{l}_{j3}$       (b) Layout II:  $c_j = o_j \vee \bar{l}_{j3} = (l_{j1} \vee \bar{l}_{j2}) \vee \bar{l}_{j3}$

**Figure EC.3** Possible layouts in  $\mathfrak{R}$  when clause node appears

Figure EC.4 shows the embedded graph  $\mathfrak{R}$  for the example in Figure 2.

**REMARK EC.1.** Note the graph  $\mathfrak{R}$  has variable, NOT, auxiliary, OR and clause vertices. Each variable vertex or auxiliary vertex is associated with a value of either a variable or its negation. Each clause vertex is associated with a value of a clause. Since the graph  $\mathfrak{R}$  is constructed from a bipartite graph that has no arc between variables, the path connecting two adjacent clause vertices is composed of vertices derived only from the same variable. For example, the path could consist of vertices  $u_i$ ,  $w_i$  or  $t_i$  as they are all derived from the same variable  $u_i \in \mathcal{U}$ . Based on our way of adding NOT and auxiliary vertices, the OR and clause vertices do not directly connect to any NOT vertices.

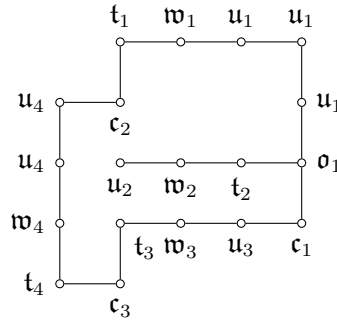


Figure EC.4 The embedded graph  $\mathfrak{R}$  for the example in Figure 2.

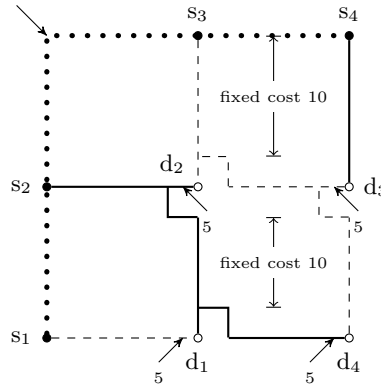
## A.2. Flow network design

It is well-known that multiechelon lot-sizing problem can be modeled as a single source network flow problem in a two-dimensional grid, see Zangwill (1969). In this section, we use a basic network in Figure EC.5 to replace variable and auxiliary vertices in the embedded graph  $\mathfrak{R}$ . With careful design on the connections, we can show that the satisfiability problem on  $\mathfrak{R}$  is related to a Muls problem. Throughout this subsection, we make the following conventions in our flow network design:

- The arcs that are not shown in the network design have high fixed costs or inventory holding costs. Thus, they are prohibited in the minimal cost flow,
- The dotted arcs are free, i.e. no cost is associated with them,
- The demand nodes are white dots in each figure with 1 unit demand,
- Vertical arcs connect the same time periods between different echelons. The fixed costs are shown in each figure and the production costs are always 0,
- Horizontal arcs connect periods in the same echelon. Most of the inventory holding costs are 0. The nonzero inventory holding cost occurs only in one period right before each demand node (with one exception in Figure EC.8b which will be elaborated later) and the value is pointed out by an arrow in figures.

**A.2.1. Basic network** The basic network is designed as in Figure EC.5 and it is the most important building block. This simple network has four potential sources  $s_1, \dots, s_4$  and four demand nodes  $d_1, \dots, d_4$ . The inventory holding cost is 5 in one period before each demand node. As shown in Figure EC.5, from the echelon of  $s_3$  to the echelon at one level higher than the echelon of  $d_2$ , we have fixed cost 10 for both dashed and solid vertical lines. Similar fixed cost structure between nodes  $d_2$  and  $d_1$  ( $d_3$  and  $d_4$ ) is demonstrated in Figure EC.5. Note that the solid and dashed lines are used in Figure EC.5 for the purpose of better illustration only, and they are not crucial in our proof. The most important thing for a basic network is its parity which is defined later in Definition EC.1.

If we connect all sources by free arcs (i.e. dotted arcs), the network turns out to have only one single source, which is indicated by the arrow at the top-left corner. So, it is an MULS and we have the following property



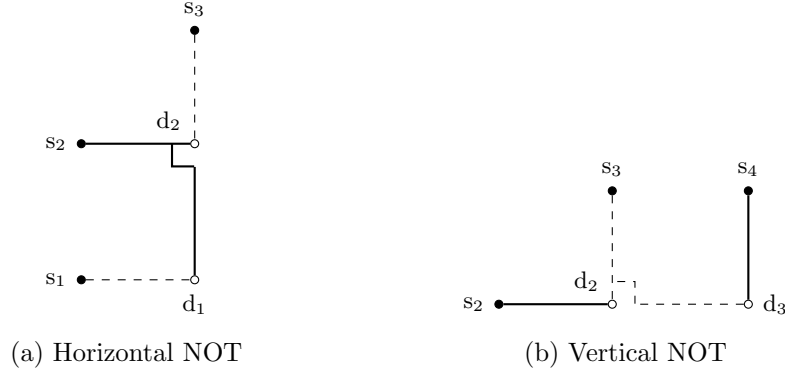
**Figure EC.5** Basic network

**PROPERTY 3** *There are only two optimal solutions to the MULS instance on the basic network with optimal cost 30 such that the demands are served through either solid arcs or dashed arcs exclusively.*

*Proof* Because we do not have production cost, the demand of each node will be served by either inventory (through horizontal arc) or transportation (through vertical arc) exclusively. Note that  $d_4$  has to be satisfied by  $s_2$  or  $s_3$ . Same does  $d_2$ . If  $d_4$  is satisfied by  $s_2$ , then  $d_1$  should be satisfied by  $s_2$  as well because the fixed cost between  $d_2$  and  $d_1$  is already charged while serving  $d_4$ .  $d_2$  should be satisfied by  $s_2$  because the inventory cost is 5 in the period before  $d_2$ , which is less than the fixed cost 10 if it is served by  $s_3$ . Thus, all four demand nodes should be served through solid arcs and the total cost 30 can be easily calculated. A similar argument can show that all four demand nodes should be served through dashed arcs if  $d_4$  is satisfied by  $s_3$ .  $\square$

**DEFINITION EC.1.** We define the parity of a basic network as 1 if the demand of  $d_4$  is fully served from the vertical arc in the solution flow, or 0 otherwise.

We design NOT networks as a half piece of the basic network shown in Figure EC.6 with either horizontal or vertical orientations. In the next subsection, we will show that this design will help us change the parity of two adjacent basic networks. So it has the same effect as NOT vertex to perform as a negation operator.



**Figure EC.6** NOT network

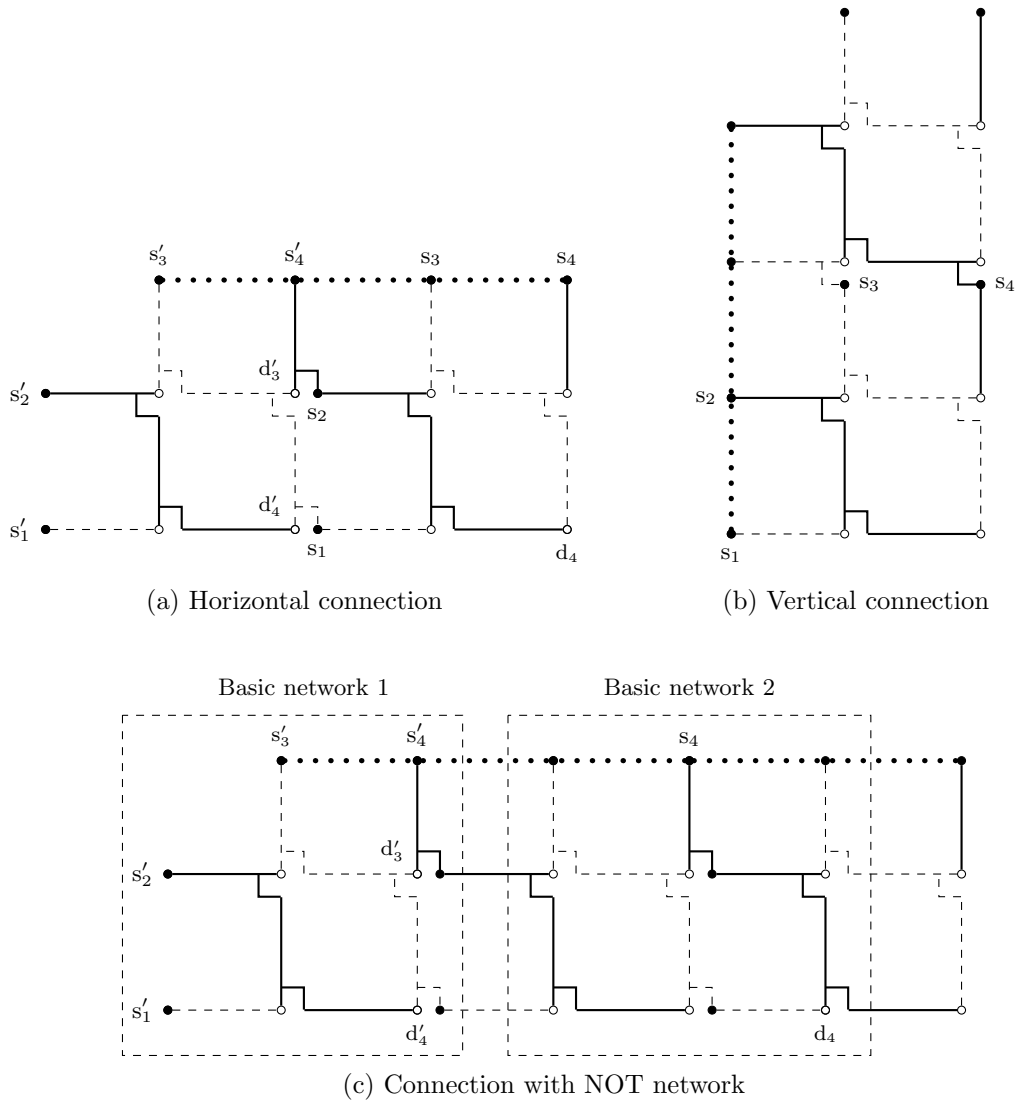
**A.2.2. Network connections** With the design of basic and NOT network, we first replace all variable and auxiliary vertices in  $\mathfrak{R}$  by basic networks, and replace NOT vertices in  $\mathfrak{R}$  by NOT networks. When vertices are horizontally connected in  $\mathfrak{R}$ , we connect the basic networks as in Figure EC.7a, otherwise we connect them as in Figure EC.7b. Since the NOT network is just a half piece of the basic network, the way of connection NOT network is shown in Figure EC.7c. Note that such design of connections could leave some sources in the basic (or NOT) network disconnected. For example, in the horizontal connection, the sources  $s_1$  and  $s_2$  are connected to the left, but the other two sources are disconnected. In this case, we connect those sources to their left sources, as in Figure EC.7a, by free arcs. The idea is that all disconnected sources can be reached freely from the single source after we have a single source network flow model built. The vertical connection is shown in Figure EC.7b.

**PROPERTY 4** *Based on the connections described above, we have the following two properties*

1. *two adjacent basic networks have the same parity;*
2. *if two basic networks are connected through a NOT network, then they have different parities.*

*Proof* In the horizontal connection, Figure EC.7a, suppose that  $d_4$  is fully served by  $s_2$ . Because of Property 3, the optimal flow is composed of solid arcs in the right basic network, i.e. the parity is 1. Apparently, the flow to  $s_2$  is coming from  $s'_4$ . Since the fixed cost is already charged,  $d'_3$  should be served by  $s'_4$  in an optimal solution. Suppose that  $d'_4$  is served by  $s'_3$ . It implies that two fixed costs with total 20 are needed. Considering that  $d'_3$  is satisfied with a fixed cost 10, we have then cost 30 without satisfying  $d'_1$  yet. Apparently, this is not an optimal solution as we know that we can serve all demand nodes with total cost 30 from Property 3. So  $d'_4$  should be served by  $s'_2$  with a total cost 30 and parity 1. Similarly, we can check the cases when  $d_4$  is fully served by  $s_3$  or in vertical connections. Therefore, the first property holds.

Note that the NOT network is just half piece of the basic network. As in Figure EC.7c, we have Basic network 1 and 2 that are connected by a NOT network. From previous discussion, we



**Figure EC.7** Connecting basic networks

know that in the optimal solution the network is connected by either solid arcs or dashed arcs. A slight difference in Basic network 2 is that the solid arcs are replaced with dashed arcs and vice versa. This different presentation of basic network is without loss of generality since it is irrelevant to the parity definition. It is easy to see that because the NOT network is just half piece of the basic network, the parities of Basic network 1 and 2 are different. Therefore, the second property is proved.  $\square$

Now, we show how to connect OR and clause vertices. Since the OR and clause vertices can only appear as in Layout I and II in Figure EC.3, we present connections for each layout as in Figure EC.8. The idea is that we introduce clause nodes with one unit of demand, denoted as  $d_5$  in Figure EC.8 such that the demands can always be satisfied (i.e., the MULS is feasible), but demands of all clause nodes are satisfied with the lowest cost if and only if the clauses are satisfied.

First, we consider the connections for Layout I, Figure EC.8a. Suppose there is a flow solution satisfies all demand nodes in all basic networks such that the flow is optimal in each basic network, i.e., the flow includes either solid arcs or dashed arcs exclusively in each basic network. Note  $d_5$  can always be satisfied through either solid arc or dashed arc. However, to achieve the minimal cost,  $d_5$  should be served through solid arc if the top basic network has parity 1 or through dashed arc if the left basic network has parity 0. Otherwise, 5 units additional inventory holding cost will be charged in node  $d_1$  or  $d_4$ . So, if the top basic network has parity 1 or the left basic network has parity 0, we can obtain the minimal additional cost of 10 (comparing to 15 that additional 5 holding cost is needed) to serve  $d_5$ . This is important in our proof because  $c_j = l_{j1} \vee \bar{l}_{j3}$  in Layout I and  $c_j$  is true only if  $l_{j1} = 1$  or  $l_{j3} = 0$ . Same reason of avoiding inventory holding costs, the optimal cost 10 of serving  $d_5$  in Figure EC.8b is obtained only if the top basic network has parity 1 or one of the left basic networks has parity 0. Note that  $d_5$  in Figure EC.8b is connected by free arc to the above, and as the only exception mentioned at the beginning of this subsection, we have 10 units of inventory holding cost in one period before node OR which has no demand. Thus, we conclude

**PROPERTY 5** *Suppose there is a flow solution satisfies all demand nodes in all basic networks such that the flow is optimal in each basic network. The clause node  $d_5$  can be satisfied with cost of 10 if*

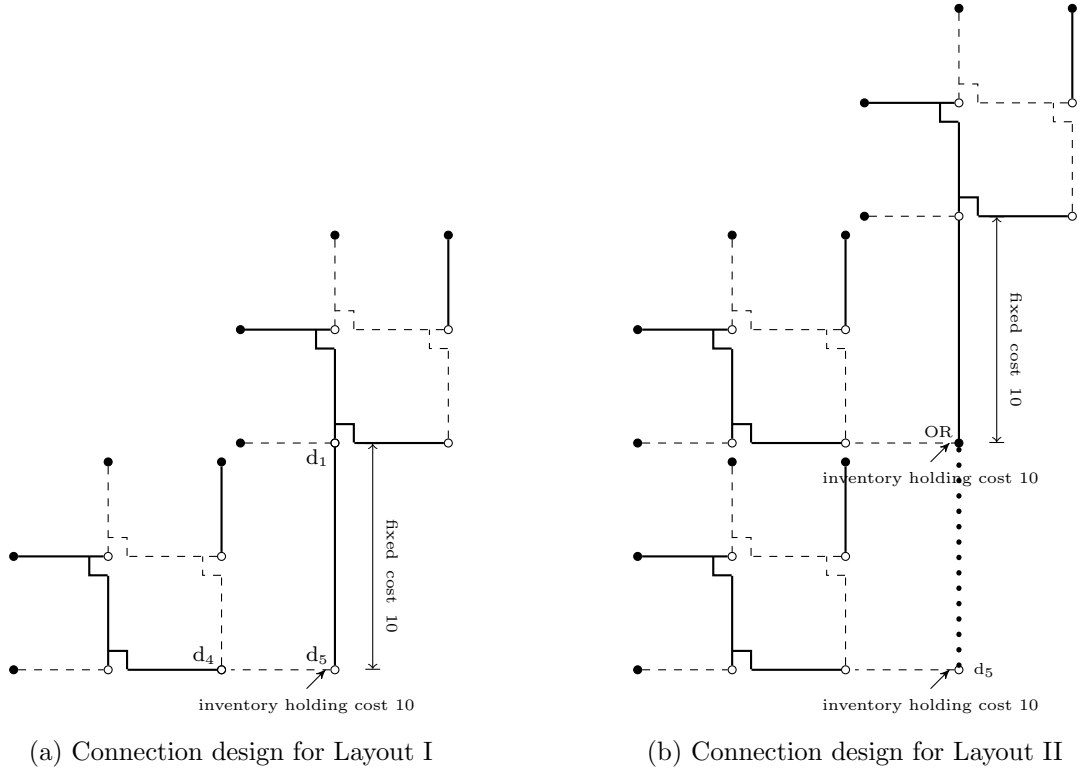
- *the top basic network has parity 1 or the left basic network has parity 0 in the case of Layout I in figure EC.3;*
- *the top basic network has parity 1 or one of the left basic networks has parity 0 in the case of Layout II in figure EC.3.*

Since all basic networks and NOT networks are connecting to either each other or the left/upper ones through free arcs, we eventually will have a network in a two-dimensional grid with a single source at the upper-left corner. Therefore, we obtain an MULS with fixed-charge production and transportation costs, and linear inventory holding costs, i.e., fixed-charge cost structure.

### A.3. Proof of Theorem 1

*Proof* For a given instance of planar 3SAT, we have showed that we can embed the identification graph  $\mathfrak{G}$  into a two-dimensional grid to have a graph  $\mathfrak{R}$ . Suppose that the graph  $\mathfrak{R}$  has  $n_1$  variable and auxiliary vertices,  $n_2$  NOT vertices and  $n_3$  clause vertices. The network design shows that the graph  $\mathfrak{R}$  can be converted to an MULS with the fixed-charge cost structure. Since the minimal costs of satisfying demands in each basic and NOT network are 30 and 15 respectively and there is additional cost of 10 for each clause node, the lowest possible cost to satisfy all demands is  $30n_1 + 15n_2 + 10n_3$ , denoted as  $\mathfrak{L}$ .

If the 3SAT is satisfiable, then for each variable  $u_i$  we let the parity of the basic network corresponding to  $u_i$  be 1 if  $u_i = 1$ , otherwise 0. From Property 4, the parities of all the basic network



**Figure EC.8** Connection design for Layouts in Figure EC.3.

will be compatible and the cost of  $30\mathbf{n}_1 + 15\mathbf{n}_2$  is obtainable to satisfy all basic and NOT networks. Since all clauses are satisfied, from Property 5, all clause nodes can be satisfied with cost of 10. Thus, the total cost  $\mathcal{L}$  can be obtained, i.e., there is a feasible solution to the MULS with total cost  $\mathcal{L}$ , which is also an optimal solution because  $\mathcal{L}$  is the lowest possible cost.

On the other hand, assume there is an optimal solution to the MULS with total cost  $\mathcal{L}$ . We will show that the 3SAT is satisfiable. Since  $\mathcal{L}$  is the lowest possible cost, we know the flow solution in each basic network includes either solid arcs or dashed arcs exclusively. Then we can assign values to the corresponding variables based on the parity. Because the clause nodes are satisfied with additional cost of 10, the parities of all basic networks imply the satisfiability of all clauses. Thus, the 3SAT is satisfiable.

Therefore, the MULS with the fixed-charge cost structure is NP-hard because the planar 3-SAT problem is NP-complete as our construction is of polynomial size.  $\square$

### Appendix B: Computing the value of $g(s_1, \beta_1; s_2, \beta_2)$

We first design a network to obtain the value of  $g((l-1, d_{l-1}^2), (m, d_m^2))$ , where  $l-1$  and  $m$  are two consecutive regeneration points at the second echelon, by calculating the shortest path between the appropriate nodes of the underlying network. In notational convenience, we denote the network as network- $(l, m)$ . Lee et al. (2003) define a transportation period as a full-track load (FTL) period

if the transportation quantity is  $W$ , otherwise it is called less-than-truck load (LTL). They show that there exists an optimal solution such that there is at most one LTL period between  $l$  and  $m$ .

Following the network design and notations by Lee et al. (2003), we know that for a given period, either all periods before are FTL periods or all periods after are FTL periods. Based on those two scenarios, we have two types of nodes as follows.

- For a period  $u$  with  $l < u < m$ , if all demands  $d^2(l, u - 1)$  at echelon 2 from period  $l$  to  $u - 1$  and  $b(u, 1, l) < d_u^2$  units of demand in period  $u$  are satisfied by FTL periods, where  $b(u, 1, l) = \lceil d^2(l, u - 1)/W \rceil W - d^2(l, u - 1)$  obviously. The node  $(u, 1, l)$  is created;
- For a period  $v$  with  $l < v < m$ , if all demands  $d^2(v + 1, m)$  at echelon 2 from period  $v + 1$  to  $m$  and  $b(v, 2, m) < d_v^2$  units of demand in period  $v$  are satisfied by FTL periods, where  $b(v, 2, m) = d^2(v, m) - \lceil d^2(v + 1, m)/W \rceil W$ . The node  $(v, 2, m)$  is created.

We also add two dummy nodes  $(l - 1, 1, l - 1)$  with  $b(l - 1, 1, l - 1) = d_{l-1}^2$  and  $(m, 2, m)$  with  $b(m, 2, m) = d_m^2$ . Each arc represents a potential transportation, we have three cases:

- Arcs  $(u, 1, l) \rightarrow (v, 1, l)$  with  $l < u < v < m$  and  $(l - 1, 1, l - 1) \rightarrow (v, 1, l)$  with  $l < u < m$ . Each arc indicates a potential FTL transportation without any preceding LTL period;
- Arcs  $(u, 2, m) \rightarrow (v, 2, m)$  with  $l < u < v \leq m$ . Each arc indicates a potential FTL transportation without any following LTL period;
- Arcs  $(u, 1, l) \rightarrow (v, 2, m)$  with  $l < u < v \leq m$  and Arcs  $(l - 1, 1, l - 1) \rightarrow (v, 2, m)$  with  $l < v \leq m$ . Each arc indicates a potential LTL transportation.

The corresponding dispatch quantity for each arc  $(u, \cdot, \cdot) \rightarrow (v, \cdot, \cdot)$  is

$$\begin{aligned} A((u, \cdot, \cdot) \rightarrow (v, \cdot, \cdot)) &= d_u^2 - b(u, \cdot, \cdot) + d^2(u + 1, v - 1) + b(v, \cdot, \cdot) \\ &= d(u, v - 1) - b(u, \cdot, \cdot) + b(v, \cdot, \cdot) \end{aligned}$$

and the cost associated with each arc is to find a transportation period  $r$  with the minimal total cost, which is

$$\begin{aligned} &p((u, \cdot, \cdot) \rightarrow (v, \cdot, \cdot)) \\ &= \min_{u \leq r \leq v} \left[ c_r^2(A((u, \cdot, \cdot) \rightarrow (v, \cdot, \cdot))) + W(u, r) - b(u, \cdot, \cdot) \cdot \sum_{t=u}^{r-1} w_t^2 \right. \\ &\quad \left. + H(r, v - 1) + b(v, \cdot, \cdot) \cdot \sum_{t=r}^{v-1} h_t^2 \right] \end{aligned} \quad (\text{EC.1})$$

where  $W(i, j)$  denotes the corresponding backlogging cost of satisfying  $d(i, j)$  at period  $j$ , and  $H(i, j)$  denotes the corresponding backlogging cost of satisfying  $d(i, j)$  at period  $i$ , which are

$$W(i, j) = \sum_{s=i+1}^j \sum_{t=i}^{s-1} w_t^2 d_s^2 \quad \text{and} \quad H(i, j) = \sum_{s=i}^{j-1} \sum_{t=s}^{j-1} h_t^2 d_s^2$$

The major difference from Lee et al. (2003) is that we do not need to consider production period in the network design, which is called replenishment and denoted as index  $k$  by Lee et al. (2003). For given  $l$  and  $m$ , the number of arcs in the network is  $O((m-l)^2)$  and it takes  $O(v-u)$  time to evaluate EC.1. So it takes  $O(T^5)$  to obtain all the arc costs of network- $(l, m)$  for all  $l$  and  $m$ . In each network, we find the shortest paths between any pair of nodes using the Floyd-Warshall method in  $O((m-l)^3)$  time. Hence, it takes  $O(T^5)$  time to find the lengths of all the shortest paths between each pair of nodes in each networks- $(l, m)$ .

As shown by Lee et al. (2003) in their Section 6.2, for given  $(s_1, \beta_1; s_2, \beta_2)$ , we can have two consecutive regeneration points  $l-1$  and  $m$  at the second echelon with  $1 \leq l \leq s_1 < s_2 \leq m \leq T$  if  $x < d_{s_1}^2$ , or  $0 \leq l-1 \leq s_1 < s_2 \leq m \leq T$  if  $x = d_{s_1}^2$ . The value of  $g(s_1, \beta_1; s_2, \beta_2)$  can be obtained as the length of the shortest path between an appropriate pair of nodes in networks- $(l, m)$ , which has been shown above and takes  $O(T^5)$  time.

Note that in (EC.1) the transportation period  $r$  will be found to minimize the cost. A question that may arise is that  $r$  is not necessarily less than  $t'$  in (10). As discussed in Remark 2, this still brings a feasible production and transportation plan. However, the solution corresponding to this feasible production and transportation plan is not extreme because the ZIO property is violated at echelon 1. Therefore, this solution will be ruled out as we look for a solution with minimal total cost, which has to be an extreme solution.

### Appendix C: Proof of Proposition 3

First, we will give two definitions in Subsection C.1. Then, the main proof of Proposition 3 is in Subsection C.2. The lemmas stated in Subsection C.2 are proved in Subsections C.3–C.6.

#### C.1. Definitions

We introduce two important definitions (with some insights provided in Remark EC.2 below) and some notations.

- A node  $(\ell_a, s_a)$  is deliverable from a node  $(\ell_b, s_b)$  if there is a flow from  $(\ell_b, s_b)$  to  $(\ell_a, s_a)$  such that the flow goes only from left to right or from top to bottom.
- Two nodes  $(i, s_a)$  and  $(i, s_b)$  are connected from above (below) if they are connected after removing all the arcs between level  $i$  and  $i+1$  ( $i-1$ ). Note that  $(i, s_a)$  and  $(i, s_b)$  are always connected from above when  $i = L$ .

Let  $(i, [s_a, s_b])$  with  $s_a \leq s_b$  be the set of nodes  $\{(i, s_a), \dots, (i, s_b)\}$  at the same level  $i$ . For notational convenience, we define that a node  $(\ell, s^\dagger)$  is deliverable from  $(\ell', [s_a, s_b])$  if  $(\ell, s^\dagger)$  is deliverable from at least one node in  $(\ell', [s_a, s_b])$ , and a node  $(\ell, s^\dagger)$  connects  $(\ell', [s_a, s_b])$  from above (below) if  $(\ell, s^\dagger)$  connects at least one node in  $(i, [s_a, s_b])$  from above (below). For example, in Figure 6, (3, 6)

is deliverable from  $(2, [1, 3])$ , but not  $(2, [2, 3])$ .

Next, we provide a few insights on the definitions.

REMARK EC.2. Suppose  $(\ell_a, s_a)$  is deliverable from a node  $(\ell_b, s_b)$ . Apparently, we must have  $\ell_b \leq \ell_a$  and  $s_b \leq s_a$ . the demand of node  $(\ell_a, s_a)$  is partially or fully served from the transportation quantity  $x_{s_b}^{\ell_b}$  only if node  $(\ell_a, s_a)$  is deliverable from node  $(\ell_b, s_b)$ . Since the subplan is a connected component, there is a path between any two nodes, but it is not necessary that one node is deliverable from another. For example, in Figure 6,  $(2, 2)$  and  $(3, 6)$  are connected, but the products cannot be shipped from node  $(2, 2)$  to serve the demand of node  $(3, 6)$ . Based on the definition,  $(3, 6)$  is deliverable from  $(2, 1)$ . However, we do not know if the demand of  $(3, 6)$  is served by the transportation from node  $(2, 1)$ , because the demand could be served by production  $x_6^1$  of node  $(1, 6)$ . So the concept of “deliverable” is to characterize the possibility of delivery products from one node to another and irrelevant to what happens physically.

The definition of the connection from below (above) is mainly used to induce contradictions by identifying cycles (after removing all the production arcs). A simple fact is that if two nodes  $(i, s_a)$  and  $(i, s_b)$  are connected from above, then they cannot be connected from below. For example, removing all the production arcs in Figure 6, nodes  $(2, 1)$  and  $(2, 6)$  are connected from below, but not above. Note that nodes  $(2, 8)$  and  $(2, 9)$  are not connected either above or below.

## C.2. Main proof

*Proof of Proposition 3* With the definitions and notations above, we will prove the proposition by assuming that Lemma EC.1–EC.4 in this subsection hold and these lemmas are proved later in Appendix C.3–C.6. Given the node  $(i, t)$  with  $i > 1$  in the subplan  $(v, w)$ , we consider the possible values of the cumulative transportation quantity from echelon  $i - 1$  to echelon  $i$  up to and including period  $t$  in the subplan, i.e.,  $X_i$ . We define  $s_M$  as the earliest period and  $s_m$  as the latest period satisfying the following two conditions:

- (c1)  $s_M > t$  and  $(i, s_M)$  is connected to some node  $(i, \tau)$  with  $\tau \leq t$  from below;
- (c2) Suppose that  $s_M$  exists.  $s_m \leq t$  and  $(i, s_m)$  is connected to  $(i, s_M)$  from below.

If  $s_M$  does not exist, then we have

LEMMA EC.1. *Suppose  $s_M$  does not exist. If  $(j_a, s_a)$  is deliverable from  $(i, [v_i + 1, t])$ , then any node  $(j_a, s_b)$  with  $s_b \leq s_a$  cannot be deliverable from  $(i, [t + 1, w_i])$ .*

Let  $\mu_\ell^3$  be the last period at each level  $\ell \in [i, L]$  such that  $(\ell, \mu_\ell^3)$  is deliverable from  $(i, [v_i + 1, t])$ . Lemma EC.1 implies that the demands of nodes in  $(\ell, [v_\ell + 1, \mu_\ell^3]) \forall \ell \in [i, L]$  are fully satisfied by  $X_i$  and we have

$$X_i = \sum_{\tau=v_i+1}^t x_\tau^i = \sum_{\ell=i}^L d^\ell(v_\ell + 1, \mu_\ell^3)$$

Therefore,  $X_i$  is in Form (3).

In the rest of the proof, we assume  $s_M$  exists and hence  $s_m$  is well defined. We give a lemma as follows.

LEMMA EC.2. *There exists a node  $(i, t')$  in  $(i, [s_m + 1, s_M])$  such that no node in  $(i, [s_m, t' - 1])$  connects any node in  $(i, [t', s_M - 1])$  from below. Let  $\alpha_\ell$  be the earliest period in each level  $\ell \in [1, i]$  such that  $(i, t')$  is deliverable from  $(\ell, \alpha_\ell)$ . We have that the period  $\alpha_\ell - 1$  is in the subplan and*

$$\sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 = \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{t'-1} x_\tau^i. \quad (\text{EC.2})$$

Based on Lemma EC.2, we consider two cases that  $t < t'$  or  $t' \leq t$ . Case (i): Suppose  $t < t'$ . We have

LEMMA EC.3. *If  $(j_a, s_a)$  is deliverable from  $(i, [t + 1, t' - 1])$ , then*

- (a) *any node  $(j_a, s_b)$  with  $s_b \leq s_a$  is not deliverable from  $(i, [t', w_i])$ , and*
- (b) *any node  $(j_a, s_b)$  with  $s_b \geq s_a$  is not deliverable from  $(i, [v_i + 1, t])$ .*

Lemma EC.3 shows that if nodes  $(\ell, \eta_\ell^1)$  and  $(\ell, \mu_\ell^1)$  are deliverable from  $(i, [t + 1, t' - 1])$  for some  $\ell \in [i, L]$  and  $\eta_\ell^1, \mu_\ell^1 \in [v_\ell + 1, w_\ell]$ , then the demands of nodes  $(\ell, \eta_\ell^1), \dots, (\ell, \mu_\ell^1)$  are fully satisfied by  $\sum_{\tau=t+1}^{t'-1} x_\tau^i$ . Apparently, if a node at level  $\ell$  is deliverable from  $(i, [t + 1, t' - 1])$ , then there exist nodes at level  $\ell - 1$  that are deliverable from  $(i, [t + 1, t' - 1])$ . Therefore, there exist  $j_1 \in [i - 1, L - 1]$ , and  $\eta_\ell^1, \mu_\ell^1 \in [v_\ell + 1, w_\ell] \forall \ell \in [i, j_1]$  such that

$$\sum_{\tau=t+1}^{t'-1} x_\tau^i = \sum_{\ell=i}^{j_1} d^\ell(\eta_\ell^1, \mu_\ell^1). \quad (\text{EC.3})$$

Note that  $j_1 < L$ . If not, then a node at the final level  $L$  is deliverable from a node, say  $(i, \tau)$ , in  $(i, [t + 1, t' - 1])$ . Hence, the delivery path will intersect with the path that connects  $(i, s_m)$  and  $(i, s_M)$  from below, which implies that  $(i, \tau)$  and  $(i, s_m)$  are connected from below. Since  $\tau < t' \leq s_M$  and satisfies condition (c1), it violates the choice of  $s_M$ . Therefore, we have  $j_1 < L$ . We allow  $j_1 = i - 1$  when  $\sum_{\tau=t+1}^{t'-1} x_\tau^i = 0$ , in case  $t' = t + 1$  and  $[t + 1, t' - 1] = \emptyset$ . Combining (EC.2) and (EC.3), we have

$$\sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 = \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{t'-1} x_\tau^i$$

$$\begin{aligned}
&= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^t x_\tau^i + \sum_{\tau=t+1}^{t'-1} x_\tau^i \\
&= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^t x_\tau^i + \sum_{\ell=i}^{j_1} d^\ell(\eta_\ell^1, \mu_\ell^1),
\end{aligned}$$

which implies

$$X_i = \sum_{\tau=v_i+1}^t x_\tau^i = \sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) - \sum_{\ell=i}^{j_1} d^\ell(\eta_\ell^1, \mu_\ell^1),$$

where  $\sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1$  is the production quantity of a sequence of periods and we can relabel  $\alpha_\ell - 1$  as  $\mu_\ell^1 \forall \ell \in [1, i-1]$ . This is exactly Form (1).

Case (ii): Suppose  $t' \leq t$ . We have

LEMMA EC.4. *If  $(j_a, s_a)$  is deliverable from  $(i, [t', t])$ , then*

- (c) *any node  $(j_a, s_b)$  with  $s_b \leq s_a$  is not deliverable from  $(i, [t+1, w_i])$ , and*
- (d) *any node  $(j_a, s_b)$  with  $s_b \geq s_a$  is not deliverable from  $(i, [v_i+1, t'-1])$ .*

Note that Lemma EC.4 is similar to Lemma EC.3. Then there exist  $j_2 \in [i-1, L-1]$ , and  $\eta_\ell^2, \mu_\ell^2 \in [v_\ell + 1, w_\ell] \forall \ell \in [i, j_2]$  such that

$$\sum_{\tau=t'}^t x_\tau^i = \sum_{\ell=i}^{j_2} d^\ell(\eta_\ell^2, \mu_\ell^2) \quad (\text{EC.4})$$

Note that  $j_2 < L$ . If not, then a node at the final level  $L$  is deliverable from a node, say  $(i, \tau)$ , in  $(i, [t', t])$ . Hence, the delivery path will intersect with the path that connecting  $(i, s_m)$  and  $(i, s_M)$  from below, which implies that  $(i, \tau)$  and  $(i, s_M)$  are connected from below. Since  $\tau \geq t' > s_m$  and satisfies condition (c2), it violates the choice of  $s_m$ . Therefore, we have  $j_2 < L$ . Combining (EC.2) and (EC.4), we have

$$\begin{aligned}
\sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 &= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{t'-1} x_\tau^i \\
&= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^t x_\tau^i - \sum_{\tau=t'}^t x_\tau^i \\
&= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^t x_\tau^i - \sum_{\ell=i}^{j_2} d^\ell(\eta_\ell^2, \mu_\ell^2),
\end{aligned}$$

which implies

$$X_i = \sum_{\tau=v_i+1}^t x_\tau^i = \sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 - \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\ell=i}^{j_2} d^\ell(\eta_\ell^2, \mu_\ell^2)$$

where  $\sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1$  is the production quantity of a sequence of periods and we can relabel  $\alpha_\ell - 1$  as  $\mu_\ell^2 \forall \ell \in [1, i-1]$ . This is exactly Form (2).

Next, we prove all the lemmas in the Proof of Proposition 3 in the following subsections.

### C.3. Proof of Lemma EC.1

We will prove the lemma by introducing contradictions. Let  $(j_a, s_a)$  be deliverable from  $(i, s_1)$  with  $s_1 \in [v_i + 1, t]$  and  $(j_a, s_b)$  be deliverable from  $(i, t_1)$  with  $t_1 > t$  as in Figure EC.9. The solid lines indicate the deliverable property. Apparently, they must intersect as shown in Figure EC.9. Since  $t_1$  satisfies condition (c1), it contradicts to the assumption that  $s_M$  does not exist.  $\square$

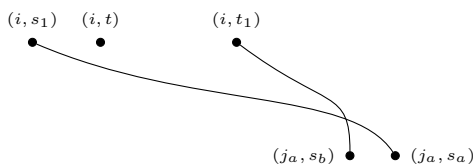


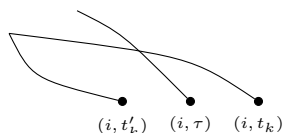
Figure EC.9 Illustration of Proof of Lemma EC.1

### C.4. Proof of Lemma EC.2

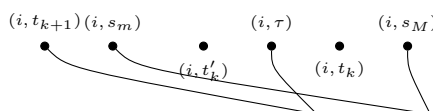
*Proof of Lemma EC.2* First we will show that  $t'$  can be found by an algorithm of repeating the following two steps. We denote  $t_0 = s_M$  to simplify the notation. At each iteration  $k$ , starting from  $k = 0$ , we

Step 1. find  $t'_k$  be the earliest period in the subplan that  $t'_k \leq t_k$  and  $(i, t'_k)$  connects to  $(i, t_k)$  from above;

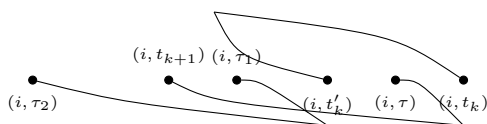
Step 2. find  $t_{k+1}$  be the earliest period in the subplan such that  $t_{k+1} < t'_k$  and  $(i, t_{k+1})$  connects to  $(i, [t'_k, t_k - 1])$  from below.



(a) Illustration of Proof of P1



(b) Illustration of Proof of P2



(c) Illustration of Proof of P5

Figure EC.10 Illustration of Proofs.

The algorithm stops when  $t'_k = t_k$  at some iteration  $k$  and we let  $t' = t_k$  when algorithm stops. Apparently, the algorithm will stop as the time period at level  $i$  in the subplan is finite. At any iteration  $k$ , we have following properties:

P1 Any two nodes in  $(i, [t'_k, t_k])$  are connected from above.

Because every node is deliverable from a production node at level 1, the path that delivery products from level 1 to any node in  $(i, [t'_k, t_k])$  will intersect with the path connecting  $(i, t'_k)$  and  $(i, t_k)$  from above, see Figure EC.10a. Thus, any node in  $(i, [t'_k, t_k])$  connects both  $(i, t'_k)$  and  $(i, t_k)$  from above, and the property is proved.

P2 If  $t'_k > s_m$ , then  $t_{k+1} > s_m$ .

We will show it by introducing contradictions. Suppose  $t_{k+1} \leq s_m$ . By Step 2 of the algorithm, there is a node  $(i, \tau)$  in  $(i, [t'_k, t_k - 1])$  connects  $(i, t_{k+1})$  from below. We know  $(i, s_m)$  and  $(i, s_M)$  are connect from below. So the path connecting  $(i, s_m)$  and  $(i, s_M)$  will intersect with the path connecting  $(i, \tau)$  and  $(i, t_{k+1})$ , see Figure EC.10b. If  $t < \tau$ ,  $\tau$  satisfies the condition (c1). Note that  $\tau < t_k \leq s_M$ . It contradicts with the choice of  $s_M$ . If  $\tau \leq t$ ,  $\tau$  satisfies the condition (c2) and  $\tau \geq t'_k > s_m$ , which contradicts with the choice of  $s_m$ . So  $t_{k+1} > s_m$ .

P3 Any node in  $(i, [t'_k, t_k])$  is connected to a node in  $(i, [t'_0, t_0 - 1])$  if  $t'_0 < t_0 (= s_M)$ .

This is true when  $k = 0$  because of the property (P1). Suppose it is true for  $k - 1$ . Because of the property (P1), any two nodes in  $(i, [t'_k, t_k])$  are connected from above. Also, node  $(i, t_k)$  connects to  $(i, [t'_{k-1}, t_{k-1} - 1])$  from below. So the property holds.

P4 If  $t_k > s_m$ , then  $t'_k > s_m$ .

If the algorithm stops at  $k = 0$ , then  $t' = t'_0 = t_0 = s_M$ . The property holds obviously. Now we assume that the algorithm does not stop at  $k = 0$ . Thus,  $t'_0 < t_0$ . Suppose the property is not true and  $t'_k \leq s_m$ , i.e., node  $(i, s_m)$  is in  $(i, [t'_k, t_k])$ . The Property (P3) shows that  $(i, s_m)$  connects to  $(i, [t'_0, t_0 - 1])$ . Together with  $(i, s_M)$  connects to  $(i, [t'_0, t_0 - 1])$  from above based on Property (P1) and  $(i, s_m)$  and  $(i, s_M)$  are connected from below, we obtain a cycle in the subplan which is a contradiction. So  $t'_k > s_m$ .

P5 No node in  $(i, [t_{k+1}, t_k - 1])$  connects to nodes in  $(i, [s_m, t_{k+1} - 1])$  from below.

Because  $t_0 = s_M > s_m$ . From Property (P2) and (P4), we know the algorithm will stop at a period later than  $s_m$ . So the interval  $[s_m, t_{k+1} - 1]$  in Property (P5) is valid. Since  $[t_{k+1}, t_k - 1] = [t_{k+1}, t'_k - 1] \cup [t'_k, t_k - 1]$ , we consider two cases for the node set  $(i, [t_{k+1}, t_k - 1])$  in Property (P5). If any node in  $(i, [t'_k, t_k - 1])$  connects to nodes in  $(i, [s_m, t_{k+1} - 1])$  from below, then it violates the choice of  $t_{k+1}$  in Step 2. If a node  $(i, \tau_1)$  in  $(i, [t_{k+1}, t'_k - 1])$  connects to a node  $(i, \tau_2)$  in  $(i, [s_m, t_{k+1} - 1])$  from below, see Figure EC.10c, the connecting path will intersect with the path that connecting  $(i, t_{k+1})$  to  $(i, \tau)$  from below, where the node  $(i, \tau)$  is in  $(i, [t'_k, t_k - 1])$ . Now,  $(i, \tau)$  can connects a node  $(i, \tau_2)$  in  $(i, [s_m, t_{k+1} - 1])$  from below. It violates the choice of  $t_{k+1}$ .

In the proof of Property (P5), we have already showed that the algorithm stops at  $t' \in [s_m + 1, s_M]$ . Since  $t' \leq t_{k+1}$ , Property (P5) implies that there is no node in  $(i, [t_{k+1}, t_k - 1])$  connects to nodes

in  $(i, [s_m, t' - 1])$  from below for any  $k$ . Note that  $\bigcup_k [t_{k+1}, t_k - 1] = [t', s_M - 1]$ . Thus, there is no node in  $(i, [s_m, t' - 1])$  connects node in  $(i, [t', s_M - 1])$  from below.

Now, with  $\alpha_\ell \forall \ell \in [1, i]$  defined in the lemma, we will prove two claims,

1.  $\alpha_i = t'$ ;
2.  $x_\tau^i = 0 \forall \tau \in [\alpha_{i-1}, \alpha_i - 1]$ ;

First, we show that  $\alpha_i = t'$ . Certainly,  $\alpha_i \leq t'$  by the definition of  $\alpha_i$  and deliverable property. Suppose  $\alpha_i < t'$ . Because  $(i, t')$  is deliverable from  $(i, \alpha_i)$ , the delivery path is on level  $i$  connecting  $(i, \alpha_i)$  to  $(i, t')$ . Thus  $(i, \alpha_i)$  and  $(i, t')$  are connected from above and the Step 1 should not stop, which is a contradiction. So,  $\alpha_i = t'$ .

Then we prove the second claim by showing that there is no arc from  $(i - 1, \tau)$  to  $(i, \tau)$  in the subplan for any  $\tau \in [\alpha_{i-1}, \alpha_i - 1]$  by introducing contradictions. Suppose  $\tau$  is the earliest period in  $[\alpha_{i-1}, \alpha_i - 1]$  such that the arc from  $(i - 1, \tau)$  to  $(i, \tau)$  is in the subplan. Since  $(i, t')$  is deliverable from  $(i - 1, \alpha_{i-1})$  and no arc from  $(i - 1, \tau')$  to  $(i, \tau')$  with  $\tau' \in [\alpha_{i-1}, \tau - 1]$  in the subplan (according to the assumption that  $\tau$  is the earliest),  $(i, t')$  is deliverable from  $(i - 1, \tau)$ . Thus,  $(i, \tau)$  and  $(i, t')$  are connected from above. This contradicts with the claim that Step 1 stops at  $t'$ . Because, for all  $\tau \in [\alpha_{i-1}, \alpha_i - 1]$ , there is no arc from  $(i - 1, \tau)$  to  $(i, \tau)$ , we have  $x_\tau^i = 0$ .

At the end, we will prove the rest of the Lemma. Apparently,  $\alpha_1 \leq \dots \leq \alpha_i$ . We know  $s_m < t'$  and node  $(i, s_m)$  must be deliverable from some production period at level 1. If node  $(i, s_m)$  is deliverable from production period later or at period  $\alpha_1$ , then  $(i, s_m)$  and  $(i, t')$  will be connected from above, which again violates the claim that Step 1 stops at  $t'$ . Since node  $(i, s_m)$  is in the subplan and deliverable from a production period before  $\alpha_1$ ,  $\alpha_1 - 1$  is a period at level 1 in the subplan. Same reason, we have  $\alpha_\ell - 1$  is a period at level  $\ell$  in the subplan for  $\ell \in [1, i]$  and

$$\begin{aligned} \sum_{\tau=v_1+1}^{\alpha_1-1} x_\tau^1 &= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{\alpha_{i-1}-1} x_\tau^i \\ &= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{\alpha_i-1} x_\tau^i \end{aligned} \quad (\text{EC.5})$$

$$= \sum_{\ell=1}^{i-1} d^\ell(v_\ell + 1, \alpha_\ell - 1) + \sum_{\tau=v_i+1}^{t'-1} x_\tau^i \quad (\text{EC.6})$$

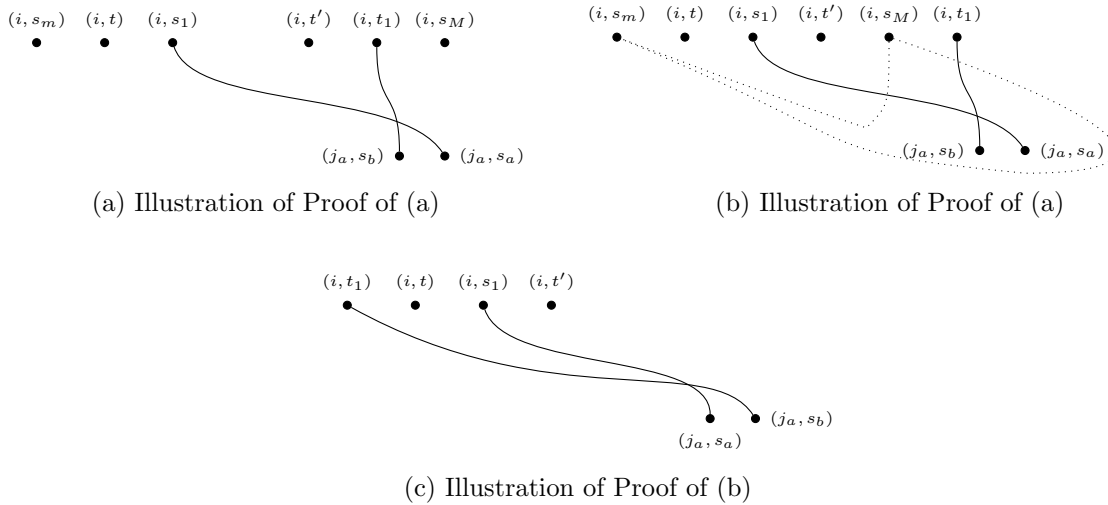
where (EC.5) holds because of claim 1 and (EC.6) holds because of claim 2.  $\square$

### C.5. Proof of Lemma EC.3

*Proof of Lemma EC.3* Let  $(j_a, s_a)$  be deliverable from  $(i, s_1)$  with  $s_1 \in [t + 1, t' - 1]$  and  $(j_a, s_b)$  be deliverable from  $(i, t_1)$ . We will prove (a) and (b) by introducing contradictions. We have nodes and deliverable paths shown in Figure EC.11.

Proof of (a): Suppose  $s_b \leq s_a$  and  $t_1 \in [t', w_i]$ . First, we suppose  $t_1 < s_M$ , see Figure EC.11a. It is obvious that those two solid paths will intersect. So  $(i, s_1)$  and  $(i, t_1)$  are connected from below, which contradicts with Lemma EC.2. Now, we suppose  $t_1 \geq s_M$ , see Figure EC.11b. Since  $(i, s_m)$  and  $(i, s_M)$  are connected from below, we use dotted lines to show two possibilities of connections. Either way we have that  $(i, s_1)$  and  $(i, s_m)$  are connected from below. Note that  $s_1 > t$ ,  $s_1 \leq t' - 1 < t' \leq s_M$  and satisfies condition (c1). So it contradicts the choice of  $s_M$ .

Proof of (b): Suppose  $s_b \geq s_a$  and  $t_1 \in [v_i + 1, t]$ . We have nodes and deliverable paths shown in Figure EC.11c. It is obvious that those two lines will intersect. Hence,  $(i, s_1)$  and  $(i, t_1)$  are connected from below. Since  $s_1 \in [t + 1, t' - 1]$  is an earlier period than  $t' \leq s_M$  and satisfies condition (c1), it contradicts the choice of  $s_M$ .  $\square$



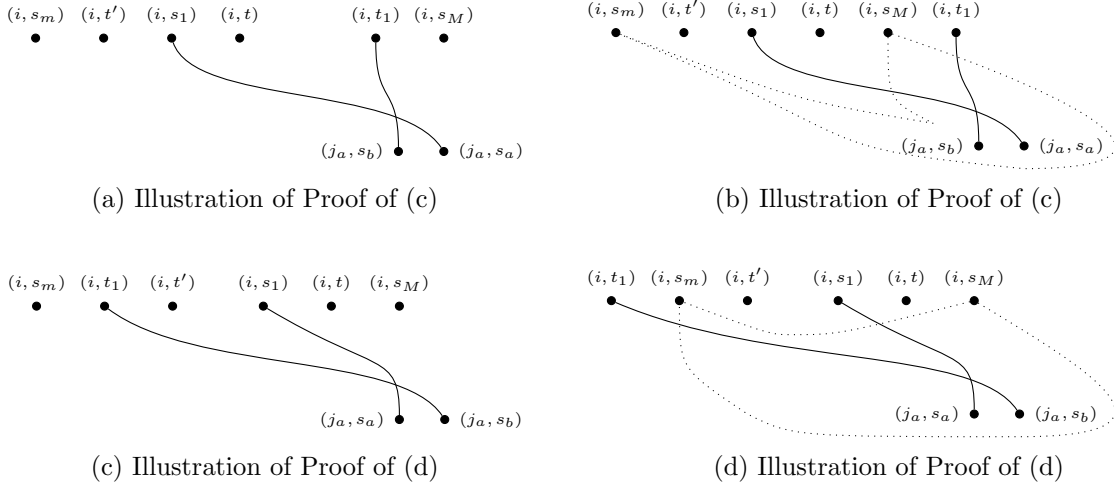
**Figure EC.11** Illustration of Proofs

## C.6. Proof of Lemma EC.4

*Proof of Lemma EC.4* Let  $(j_a, s_a)$  be deliverable from  $(i, s_1)$  with  $s_1 \in [t', t]$  and  $(j_a, s_b)$  be deliverable from  $(i, t_1)$ . We will prove (c) and (d) by introducing contradictions. We have nodes and deliverable paths shown in Figure EC.12.

Proof of (c): Suppose  $s_b \leq s_a$  and  $t_1 \in [t + 1, w_i]$ . Those two solid paths must intersect in Figure EC.12a and EC.12b because of the deliverable property. If  $t_1 < s_M$  as in Figure EC.12a, then  $t_1 > t$  is an earlier period than  $s_M$  satisfies the condition (c1), which brings a contradiction to the choice of  $s_M$ . Suppose  $t_1 \geq s_M$  as in Figure EC.12b. Since  $(i, s_m)$  and  $(i, s_M)$  are connected from below, we use dotted lines to show two possibilities of connections. Either way we have that  $(i, s_1)$  and  $(i, s_M)$  are connected from below. Since  $s_1 \geq t' > s_m$  is a later period than  $s_m$  and satisfies condition (c2), it contradicts to the choice of  $s_m$ .

Proof of (d): Suppose  $s_b \geq s_a$  and  $t_1 \in [v_i + 1, t' - 1]$ . Those two solid paths must intersect in Figure EC.12c and EC.12d because of the deliverable property. First, we suppose  $t_1 \geq s_m$ , see Figure EC.12c. Since  $(i, t_1)$  and  $(i, s_1)$  are connected from below and  $s_1 \leq t < s_M$ , it contradicts with Lemma EC.2. Now we suppose  $t_1 < s_m$ , see Figure EC.12d. Since  $(i, s_m)$  and  $(i, s_M)$  are connected from below, we use dotted lines to show two possibilities of connections. Either way we have that  $(i, s_1)$  and  $(i, s_M)$  are connected from below. Since  $s_1 \geq t' > s_m$ , it contradicts to the choice of  $s_m$ .



**Figure EC.12** Illustration of Proofs

□

#### Appendix D: Proof of Proposition 4

*Proof* Recall that  $X$  and  $\bar{X}$  are two cumulative quantity vectors up to and including periods  $t$ , and  $t-1$  respectively. We are going to show that the number of allowable values for the pair  $(X, \bar{X})$  is in  $O(T^{2L_1(L-1)+1})$ . As shown previously, the number of allowable values for the pair  $(X_1, \bar{X}_1)$  is  $O(T)$  because  $X_1 - \bar{X}_1 \in \{0, \epsilon, \mathbb{C}\}$ . Now, we consider the pair  $(X_i, \bar{X}_i) \forall i \in [2, L]$ . Given an arbitrary  $i \in [2, L]$  in the rest of proof, the goal is to prove the following statement:

(S1) there exists  $j \in [i, L]$  such that the number of allowable values of  $X_i, \dots, X_j$  and  $\bar{X}_i, \dots, \bar{X}_j$  is in  $O(T^{2L_1(j-i+1)})$ .

To understand the complexity more easily, note that we will get the same complexity if the number of allowable values for each pair  $(X_\ell, \bar{X}_\ell)$  is in  $O(T^{2L_1})$  with  $\ell \in [i, j]$ . Therefore, the statement (S1) implies that the number of allowable values in  $\Theta_{t, (\bar{v}, \bar{w})}^2$  is  $O(T^{2L_1(L-1)+1})$  and the proposition is proved if (S1) holds.

First we assume that no arc from node  $(\ell, t-1)$  to  $(\ell, t) \forall \ell \in [i, L]$  is in the subplan. Then it is easy to see that

$$\bar{X}_i = \sum_{\ell=i}^L d^\ell(\bar{v}_\ell + 1, t-1).$$

Since the number of allowable values of  $X_i$  is in  $O(T^{2L_1})$  and  $\bar{X}_i$  is given as above, the statement (S1) is proved with  $j = i$ .

Now, we can assume that there exist arcs from node  $(\ell, t-1)$  to  $(\ell, t) \forall \ell \in [i, L]$  in the subplan. Let  $j \geq i$  be the highest level that the arc from node  $(j, t-1)$  to  $(j, t)$  is in the subplan. Thus, there is no arc between  $(\ell, t-1)$  and  $(\ell, t) \forall \ell \in [i, j-1]$ . We have

$$\bar{X}_{\ell+1} = \bar{X}_\ell - d^\ell(\bar{v}_\ell + 1, t-1) \quad \forall \ell \in [i, j-1]$$

which implies that

$$\bar{X}_\ell = \bar{X}_i - \sum_{\ell=i}^{\ell-1} d^\ell(\bar{v}_\ell + 1, t-1) \quad \forall \ell \in [i+1, j]. \quad (\text{EC.7})$$

We have the following two claims.

Claim 1:  $\bar{X}_\ell$  is uniquely determined  $\forall \ell \in [i+1, j]$  when  $\bar{X}_i$  is given.

Claim 2: The number of allowable values of the pair  $(\bar{X}_\ell, X_\ell)$  is in  $O(T^{2L_1})$  as well for each  $\ell \in [i+1, j]$  when the value of  $\bar{X}_i$  is given.

Claim 1 follows (EC.7) directly. Claim 2 holds because Claim 1 and the number of allowable values of  $X_\ell$  is in  $O(T^{2L_1})$  from Remark 4. Next, we can consider two cases to conclude the proof.

Case (a): Suppose there exists  $j' \in [i+1, j] \neq \emptyset$  such that the arc from node  $(j'-1, t)$  to  $(j', t)$  is not in the subplan. Then  $X_{j'} = \bar{X}_{j'}$  and  $X_{j'}$  is also uniquely determined with given  $\bar{X}_i$  by Claim 1. Because the number of allowable values of the pair  $(\bar{X}_i, X_i)$  is in  $O(T^{4L_1})$  from Remark 4 and Claim 2, the number of allowable values of  $X_i, \dots, X_{j'}$  and  $\bar{X}_i, \dots, \bar{X}_{j'}$  is in  $O(T^{2L_1(j'-i+1)})$  and the statement (S1) is proved with  $j = j'$ .

Case (b): Suppose that either  $j = i$ , i.e.,  $[i+1, j] \neq \emptyset$ , or the arcs from node  $(\ell-1, t)$  to  $(\ell, t)$  for all  $\ell \in [i+1, j]$  are in the subplan. Recall the definition introduced in the proof of Proposition 3. We have the network as in Figure EC.13a, where  $s'_m$  is the latest period at level  $i$  that  $(j, t-1)$  is deliverable from  $(i, s'_m)$ . Note that when  $j = i$ , node  $(i, s'_m)$  connects to  $(i, t)$  through a straight line. Recall that we proved Proposition 3 by analyzing the possible values of  $X_i$ . Now, we use a similar proof on  $\bar{X}_i$ , i.e., period  $t-1$  (i.e., we substitute  $t$  to  $t-1$  in that proof). Then, we have  $s_M = t$  and  $s'_m \leq s_m$  by the definition of  $s_M$  and  $s_m$ . Since  $s_m \in [s'_m, t-1]$  and  $s_M = t$ , the  $t'$  in Lemma EC.2 must satisfy  $t' \leq t-1$ . So we have Case (ii) as in the proof for Proposition 3 and  $\bar{X}_i$  follows the Form (2).

Then, we show that the number of allowable values of  $\bar{X}_i$  is in  $O(T^{1+|\mathcal{L}_1 \cap [1, j-1]|})$ . We claim that no node on or below the solid path in Figure EC.13a is deliverable from  $(i, [s'_m + 1, t-1])$ . If it



**Figure EC.13** Illustration of the proof

is not true, then a node  $(\ell_a, s_a)$  on or below the solid path in Figure EC.13b is deliverable from  $(i, s_b)$  with  $s_b \in [s'_m + 1, t - 1]$ . Because there is no arc between  $(\ell, t - 1)$  and  $(\ell, t) \forall \ell \in [i, j - 1]$ , the path connecting  $(\ell_a, s_a)$  and  $(i, s_b)$  will intersect with the path connecting  $(i, s'_m)$  to  $(j, t - 1)$ . Then,  $(j, t - 1)$  is deliverable from  $(i, s_b)$ , which contradicts with the choice of  $s'_m$ . So the claim is correct and no demand of node at or below level  $j$  is served from node  $(i, [s'_m + 1, t - 1])$ . Therefore,  $\sum_{\tau=t'}^{t-1} x_\tau^i$  in (EC.4) (we modify the index  $t$  in (EC.4) to  $t - 1$  while replicating the same proof on  $\bar{X}_i$ ) do not include demands at level  $j$  or any level lower than level  $j$ , i.e.,  $j_2 < j$ . Hence, the number of allowable values of  $\bar{X}_i$  is in  $O(T^{1+|\mathcal{L}_1 \cap [1, j-1]|})$ .

Suppose  $j > i$ . Note that the number of allowable values of  $X_j$  and  $X_i$  are in  $O(T^{L_1+|\mathcal{L}_1 \cap [j, L-1]|})$  and  $O(T^{2L_1})$  respectively from Remark 4.  $\bar{X}_j$  is uniquely determined after  $\bar{X}_i$  because of Claim 1. The number of allowable values of  $X_i, X_j, \bar{X}_i, \bar{X}_j$  is in  $O(T^{4L_1})$  since  $1 + |\mathcal{L}_1 \cap [1, j - 1]| + |\mathcal{L}_1 \cap [j, L - 1]| \leq L_1$ . Based on Claim 2, the number of allowable values of the pairs  $(\bar{X}_\ell, X_\ell)$  are in  $O(T^{2L_1}) \forall \ell \in [i + 1, j - 1]$  when the value of  $\bar{X}_i$  is given. Hence the statement (S1) is proved.

Now we consider  $j = i$ . Since the number of allowable values of  $X_i$  is in  $O(T^{L_1+|\mathcal{L}_1 \cap [j, L-1]|})$  by Remark 4, The number of allowable values of  $X_i, \bar{X}_i$  is in  $O(T^{2L_1})$  and the statement (S1) holds with  $j = i$ .

□

---

## References

- Aggarwal A, Park J (1993) Improved algorithms for economic lot size problems. *Operations Research* 41(3):549–571.
- Ahmed S, He Q, Li S, Nemhauser G (2016) On the computational complexity of minimum-concave-cost flow in a two-dimensional grid. Accepted by *SIAM Journal on Optimization* .
- Ahuja R, Magnanti T, Orlin J (1993) *Network Flows: Theory, Algorithm, and Applications* (Englewood Cliffs, NJ: Prentice Hall).
- Atamtürk A, Küçükyavuz S (2005) Lot sizing with inventory bounds and fixed costs: Polyhedral study and computation. *Operations Research* 53(4):711–730.
- Atamtürk A, Küçükyavuz S (2008) An  $\mathcal{O}(n^2)$  algorithm for lot sizing with inventory bounds and fixed costs. *Operations Research Letters* 36(3):297–299.
- Bitran GR, Yanasse HH (1982) Computational complexity of the capacitated lot size problem. *Management Science* 28(10):1174–1186.
- Federgruen A, Tzur M (1991) A Simple Forward Algorithm to Solve General Dynamic Lot Sizing Models with  $n$  Periods in  $\mathcal{O}(n \log n)$  or  $\mathcal{O}(n)$  Time. *Management Science* 37(8):909–925.
- Florian M, Klein M (1971) Deterministic production planning with concave costs and capacity constraints. *Management Science* 18(1):12–20.
- Florian M, Lenstra JK, Rinnooy Kan AHG (1980) Deterministic production planning: Algorithms and complexity. *Management Science* 26(7):669–679.
- He Q, Ahmed S, Nemhauser G (2015) Minimum concave cost flow over a grid network. *Mathematical Programming* 150(1):79–98.
- Hwang H (2010) Economic lot-sizing for integrated production and transportation. *Operations Research* 58(2):428–444.
- Hwang H, Ahn H, Kaminsky P (2013) Basis paths and a polynomial algorithm for the multistage production-capacitated lot-sizing problem. *Operations Research* 61(2):469–482.
- Kaminsky P, Simchi-Levi D (2003) Production and distribution lot sizing in a two stage supply chain. *IIE Transactions* 35(11):1065–1075.
- Lee C, Çetinkaya S, Jaruphongsa W (2003) A dynamic model for inventory lot sizing and outbound shipment scheduling at a third-party warehouse. *Operations Research* 51(5):735–747.
- Lichtenstein D (1982) Planar formulae and their uses. *SIAM J. Comput.* 11(2):329–343.
- Love S (1972) A facilities in series inventory model with nested schedules. *Management Science* 18(5):327–338.
- Melo R, Wolsey L (2010) Uncapacitated two-level lot-sizing. *Operations Research Letters* 38:241–245.

- Niranjan S, Ciarallo F (2011) Supply performance in multi-echelon inventory systems with intermediate product demand: A perspective on allocation. *Decision Science* 42(3):575–617.
- Pochet Y, Wolsey L (2006) *Production planning by mixed integer programming* (Springer Verlag).
- Pochet Y, Wolsey LA (1993) Lot-sizing with constant batches: Formulation and valid inequalities. *Mathematics of Operations Research* 18(4):767–785.
- Shi W, Su C (2005) The rectilinear steiner arborescence problem is NP-complete. *SIAM J. Comput.* 35(3):729–740.
- Simchi-Levi D, Schmidt W, Wei Y, PY Z, Combs K, Ge Y, Gusikhin O, Sanders M, Zhang D (2015) Identifying risks and mitigating disruptions in the automotive supply chain. *Interfaces* 45(5):375–390.
- van Hoesel C, Wagelmans A (1996) An  $O(T^3)$  algorithm for the economic lot-sizing problem with constant capacities. *Management Science* 42(1):142–150.
- van Hoesel S, Romeijn HE, Morales DR, Wagelmans A (2005) Integrated lot sizing in serial supply chains with production capacities. *Management Science* 51(11):1706–1719.
- Van Vyve M (2007) Algorithms for single-item lot-sizing problems with constant batch size. *Mathematics of Operations Research* 32(3):594–613.
- Veinott AF (1969) Minimum concave-cost solution of Leontief substitution models of multi-facility inventory systems. *Operations Research* 17(2):262–291.
- Wagelmans A, Van Hoesel S, Kolen A (1992) Economic lot sizing: An  $\mathcal{O}(n \log n)$  algorithm that runs in linear time in the Wagner-Whitin case. *Operations Research* 40(1):145–156.
- Wagner H, Whitin T (1958) Dynamic version of the economic lot size problem. *Management Science* 5(1):89–96.
- Zangwill W (1968) Minimum concave cost flows in certain networks. *Management Science* 14(7):429–450.
- Zangwill W (1969) A backloging model and a multi-echelon model of a dynamic economic lot size production system - A network approach. *Management Science* 15(9):506–527.
- Zhang M, Küçükyavuz S, Yaman H (2012) A polyhedral study of multiechelon lot sizing with intermediate demands. *Operations Research* 60(4):918–935.